Analysis of flight delays in aviation system using different classification algorithms and feature selection methods

A. B. A. Anderson
Department of CSE
Veerammal Engineering College
Dindigul, TN, India

A. J. Sanjeev Kumar
Department of MCA
Thiagarajar College of Engineering, Madurai, TN, India

A. B. Arockia Christopher[iD]
abachristo123@gmail.com
Department of IT, Dr. Mahalingam College of Engineering & Technology
Pollachi, TN, India

ABSTRACT

Data mining is a process of finding correlations and collecting and analysing a huge amount of data in a database to discover patterns or relationships. Flight delay creates significant problems in the present aviation system. Data mining techniques are desired for analysing the performance in which micro-level causes propagate to make system-level patterns of delay. Analysing flight delays is very difficult – both when looking from a historical view as well as when estimating delays with forecast demand. This paper proposes using Decision Tree (DT), Support Vector Machine (SVM), Naive Bayesian (NB), K-nearest neighbour (KNN) and Artificial Neural Network (ANN) to study and analyse delays among aircrafts. The performance of different data mining methods is found in the different regions of the updated datasets on these classifiers. Finally, the result shows a significant variation in the performance of different data mining methods and feature selection for this problem. This paper aims to deal with how data mining techniques can be used to understand difficult aircraft system delays in aviation. Our aim is to develop a classification model for studying and reducing delay using different data mining methods and, in this manner, to show that DT has a greater classification accuracy. The different feature selectors are used in this study in order to reduce the number of initial attributes. Our results clearly demonstrate the value of DT for analysing and visualising how system-level effects happen from subsystem-level causes.

Keywords: Data mining; Flight delay; Classification techniques; Feature selection; Aviation

1.0 INTRODUCTION

Data mining, the extraction of hidden predictive information from large databases, is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses. Data mining tools predict future trends and behaviours, allowing businesses to make proactive, knowledge-driven decisions. The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems. Data mining tools can answer business questions that traditionally were too time-consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations. Most companies already collect and refine massive quantities of data. Data mining techniques can be implemented rapidly on existing software and hardware platforms to enhance the value of existing information resources and can be integrated with new products and systems as they are brought online.

Data mining is the automated process of analysing large volumes of data and extracting patterns in the data. Data mining tools are capable of predicting behaviours and future trends, allowing an organisation to benefit from past experience in making knowledge-driven decisions. Flight delays are obviously frustrating to air travellers and costly to airlines. Airline companies are the most important customers of the airport. A well-known phrase – 'the airplane earns only when flying' – holds true. On-time performance of the airline's schedule is a key factor in maintaining current customer satisfaction and attracting new ones. The flight schedule of the airport is the key to planning and executing an airline's operations. With each schedule, the airline defines its daily operations and commits its resources to satisfying its customers' air travel needs. Therefore, one of the basic requirements that all airlines have is to ensure a high efficiency of its ground handling activities and thus avoiding delays.

Flight delay is complex to explain. A flight can be behind schedule due to problems at the airport of origin, at the destination airport, or while it is airborne. A combination of these factors often occurs. Delays can sometimes also be attributed to the airlines themselves. Some flights are affected by reactionary delays, due to the late arrival of a previous flight. These reactionary delays can be aggravated by the schedule operation. Flight schedules are often subjected to irregularity. Due to tight connections among airlines resources, delays could dramatically propagate over time and space unless the proper recovery actions are taken. Despite the complexity, there exist some patterns of flight delay due to the schedule performance and airline itself. Some results have been extracted from the case study.

The airlines report the causes of delay in broad categories that were created by the Air Carrier On-Time Reporting Advisory Committee. The categories are Air Carrier, National Aviation System, Weather, Late-Arriving Aircraft and Security. The causes of cancellation are the same, except there is no late-arriving aircraft category.

Air Carrier: The cause of the cancellation or delay was due to circumstances within the airline's control (e.g. maintenance or crew problems, aircraft cleaning, baggage loading, fuelling, etc.).

Extreme Weather: Significant meteorological conditions (actual or forecasted) that, in the judgment of the carrier, delays or prevents the operation of a flight such as tornado, blizzard or hurricane.

National Aviation System (NAS): Delays and cancellations attributable to NAS that refer to a broad set of conditions, such as non-extreme weather conditions, airport operations, heavy traffic volume and air traffic control.

Late-Arriving Aircraft: A previous flight with the same aircraft arrived late, causing the present flight to depart late.

Security: Delays or cancellations caused by evacuation of a terminal or concourse, re-boarding of the aircraft because of a security breach, inoperative screening equipment and/or long lines in excess of 29 minutes at screening areas.

Flight delays are present every day in every part of the world. There can be flight delays due to weather, to excessive traffic, to runway construction work and to other factors, but most delays are due to weather that people assume is part of flying. But what if we could accurately predict, at least with ~70% accuracy, if a flight was going to be delayed due to weather within 10 days of the flight date? If possible, it would save passengers a great deal of time and money because passengers would be able to book flight connections with enough time to spare. By searching thoroughly through the Internet, we were able to collect a list of variables that will help create a flight delay scenario due to weather forecaster. First Waikato Environment for Knowledge Analysis (WEKA) will be used to build the model by trying different classifiers and selecting the one with the best results. WEKA is going to be used for its ease of use in accessing different classifiers and testing different settings. The flight variables used for the model have to be available at the time of ticket purchase, and at the same time, the weather variables will have to be obtained from a weather website. An early thought is that Naive Bayes (NB) will provide the best results.

In this paper, Section 2 provides a background review of airspace and delay performance metrics. Section 3 describes the algorithms used for model development. Section 4 describes the aircraft delay data sources. Section 5 presents experimental results from using WEKA methods and algorithms for delay estimation at the international level. Finally, conclusions are provided in Section 6.

2.0 BACKGROUND

Flight delay is a complex phenomenon because it can be due to problems at the origin airport, at the destination airport, or during flight. A combination of these factors often occurs. Delays can sometimes also be attributable to airlines. Some flights are affected by reactionary delays, due to late arrival of previous flights. These reactionary delays can be aggravated by the schedule operation. Flight schedules are often subjected to irregularity. Due to the tight connection among airlines' resources, delays could dramatically propagate over time and space unless the proper recovery actions are taken. Despite this complexity, flight delays are currently measurable, and there exists some pattern of flight delay due to the schedule performance and airline itself (Wu, 2005)[1].

In the United States, two government agencies keep air traffic delay statistics. The Bureau of Transportation Statistics (BTS) compiles delay data for the benefit of passengers. They define a delayed flight when the aircraft fails to release its parking brake less than 15 minutes after the scheduled departure time. The Federal Aviation Authority (FAA) is more interested in delays indicating surface movement inefficiencies and will record a delay when an aircraft requires 15 minutes or longer over the standard taxi-out or taxi-in time (Eric and Chatterji, 2002)[2].

Generally, flight delays are the responsibility of the airline. Each airline has a certain number of hourly arrivals and departures allotted per airport. If the airline is not able to get all of its scheduled flights in or out each hour, then representatives of the airline will determine

which flights to delay and which flights to cancel. These delays take one of three forms: ground delay programs, ground stops and general airport delays. When the arrival demand of an airport is greater than the determined capacity of the airport, then a ground delay program may be instituted. The airport capacity is unique to each airport, given the same weather conditions. The various facilities at an airport can determine how much traffic an airport can handle during any given weather event. Generally, ground delay programs are issued when inclement weather is expected to last for a significant period of time. These programs limit the number of aircrafts that can land at an affected airport. Because when demand is greater than the aircraft arrival capacity, flight delays will result.

Second, ground stops are issued when inclement weather is expected for a short period of time or the weather at the airport is unacceptable for landing. Ground stops mean that traffic destined to the affected airport is not allowed to leave for a certain period of time. Lastly, there are general arrival and departure delays. This usually indicates that arrival traffic is doing airborne holding or departing traffic is experiencing longer than normal taxi times or holding at the gate. These could be due to a number of reasons, including thunderstorms in the area, a high departure demand or a runway change. Our research finds that arrival and departure delays are highly correlated. Correlation between arrival and departure delays is extremely high, around 0.9, for 2002 and 2003. This finding is useful to prove that congestion at the destination airport is to a great extent originated at the departure airport.

In order to understand flight delay, it is useful to consider the phenomenon of scheduled delay. The simplest way of reducing delays is not to increase the speed and efficiency of the system to meet the scheduled time, but to push back the scheduled time to absorb the system delays. As a result, one estimate put the number of scheduled delays that were built into airline schedules in 1999 at 22.5 million minutes. The number of arrival delays reported by BTS would have been nearly 25% higher in 1999 if airlines had maintained their 1988 schedules (Wu, 2005)[1]. Sources of airport delay include many elements, such as weather, airport congestion, luggage loading, connecting passengers, etc. Weather is the main contributor to delays in the air traffic control (ATC) system. Traffic volume delays are caused by an arrival/departure demand that exceeds the normal airport arrival rate (AAR)/airport departure rate (ADR). The demand may also exceed the airport capacity if AAR and ADR are reduced due to weather conditions at the airport, equipment failure or runway closure. Delays may also be attributed to airline operation procedures (Aisling and Reynolds-Feighan, 1999)[3].

Delays to airline schedules may result from many different causes. Some are due to airport capacity limits, while some maybe due to disrupting events, e.g. missing check-in passengers at airports. According to delay analyses carried out by Eurocontrol, around 47% of delays are due to airline-related operations at airports such as aircraft turnaround operations, while the remaining delays are due to other causes such as air traffic control, weather and airport capacity constraints (Eurocontrol, 2001)[4]. The visualisation of both our algorithmic results and the statistical analyses makes it possible to discern meaningful trends in a large and complicated dataset (Phillips and Steele, 2009)[5].

The delays are based on flight data from the Enhanced Traffic Management System (ETMS) and other information sources (e.g., airline schedules, operations and delays, weather information, runway information, etc.). Aviation System Performance Metrics (ASPM) delays are a measure of actual delays experienced by the airlines and its customers. ASPM collects data at a finer granularity, reports delays of one minute or more and classifies delay by all phases of flight and time of the day. The ASPM also provides the daily number of flights cancelled by the airlines. Both systems contain data entered by human operators and are prone to data recording errors. The two systems are overlapping in certain areas and complementary in

others. Both databases can be used independently for developing NAS metrics models based on statistical analysis (Wang, Klein and Jehlen, 2009)[6]. Due to their cost and the environmental and noise issues associated with construction, it is unlikely that any new airports will be built in the near future. Therefore, to make the National Airspace System run more efficiently, techniques to more effectively use the limited airport capacity must be developed (Smith and Sherry, 2008)[7].

Correlation-based Feature Selection (CFS) subset evaluator, consistency subset evaluator, gain ratio feature evaluator, information gain attribute evaluator, OneR attribute evaluator, principal components attribute transformer, ReliefF attribute evaluators and symmetrical uncertainty attribute evaluator are used in this study in order to reduce the number of initial attributes. The classification algorithms such as Decision Tree (DT), K-nearest neighbour (KNN), Support Vector Machine (SVM), Artificial Neural Network (ANN) and NB are used to predict the warning level of the component as the classification attribution. We have explored the use of different classification techniques on aviation components data (Christopher and Balamurugan, 2014)[8].

3.0 PROPOSED METHOD

Flight delays have a weak effect on airlines, airports and passengers. Moreover, the development of accurate prediction models for flight delays became cumbersome due to the complexity of the air transportation system, the amount of methods for prediction and the deluge of data related to such system. In this context, this paper presents a thorough literature review of approaches used to build flight delay prediction models from the aircraft data perspective. We propose different data mining algorithms and summarise the initiatives used to address the flight delay prediction problem, according to scope, data and computational methods, giving special attention to an increasing usage of machine learning methods. Apart from this, we also present a timeline of major works that depict relationships between flight delay prediction problems and the research trends to address them.

In this paper, we are using different data mining methods for aircraft analysis data. The number of classifiers are NB, Functions, Lazy, Meta, MISC, Rules and DT, and the feature attribute selectors are CFS, Chi-square, Consistency (CS), Gain Ratio (GR), Information Gain, OneR, Principal Component Analysis (PCA), ReliefF (RF), Symmetrical Uncertainty (SU) and SVM attribute evaluator. These are the techniques used when building the model on training data. Then we reduce the attributes and will get the performance of the different classifiers and different feature selectors.

3.1 Feature selection

Feature selection is applied to inputs, predictable attributes or to states in a column. When scoring for feature selection is complete, only the attributes and states that the algorithm selects are included in the model-building process and can be used for prediction. If you choose a predictable attribute that does not meet the threshold for feature selection, the attribute can still be used for prediction, but the predictions will be based solely on the global statistics that exist in the model. A feature selection algorithm can be seen as the combination of a search technique for proposing new feature subsets, along with an evaluation measure that scores the different feature subsets. The simplest algorithm is to test each possible subset of features finding the one that minimises the error rate.

3.2 Classification Decision Tree (DT)

DT induction is a very popular and practical approach for pattern classification. Generally, DT is constructed in a greedy, top-down recursive manner. The tree can be constructed in a breadth first manner or depth first manner. The DT structure consists of a root node, internal nodes and leaf nodes. The classification rules are derived from the DT in the form of – 'if then else'. These rules are used to classify the records with unknown value for class label. The DT is constructed in two phases: Building Phase and Pruning Phase.

In the Building Phase of the tree, the best attributes are selected based on attribute selection measures such as Information Gain, Gain Ratio, Gini Index, etc. Once the best attribute is selected, then the tree is constructed with that node as the root node and the distinct values of the attribute are denoted as branches. The process of selecting the best attribute and representing the distinct values as branches are repeated until all the instances in the training set belong to the same class label.

In the Pruning Phase, the sub-trees that may over fit the data are eliminated. This enhances the accuracy of a classification tree. DTs handle continuous and discrete attributes. DTs are widely used because they provide human-readable rules and are easy to understand, and the construction of a DT is fast and yields better accuracy. There are several algorithms to classify the data using DTs.

The common procedure used to resolve a connection between a weather forecast and airport facility was:

- Collect data from the various available data sources,
- Use WEKA tools, format the data into a usable layout,
- Use a number of classification algorithms to connect the datasets,
- Use feature selection attributes to connect the updated datasets, and
- Test the data to ensure there is accuracy and efficiency.

The following steps are involved in the data mining methods applied to the aircraft data analysis problem:

Step 1: Collect the aircraft data from the various available data sources
Step 2: Format the data into a usable layout
Step 3: Use classification algorithms to connect the datasets
Step 4: Apply feature selection attributes to connect the updated datasets
Step 5: Analyze performance of accuracy and efficiency

In this model, we found that the feature selection method attribute reduces the number of redundant and irrelevant attributes, thereby increasing the performance of classification methods. Figure 1 displays the block diagram of proposed classification model for training samples. In this diagram, we can eliminate irrelevant features and evaluate all classifiers, thereby increasing the performance of classification accuracy and efficiency of this model. All classifier techniques and the different features attribute selection evaluator methods applied in this model for large scale aircraft dataset analysis.

The algorithm for these processes is outlined as follows:

Step 1: Input training dataset.
Step 2: • Create replica sets of the same by random selection of training examples from the dataset.
       • Learn the classifier by the different algorithms based on the actual training set.
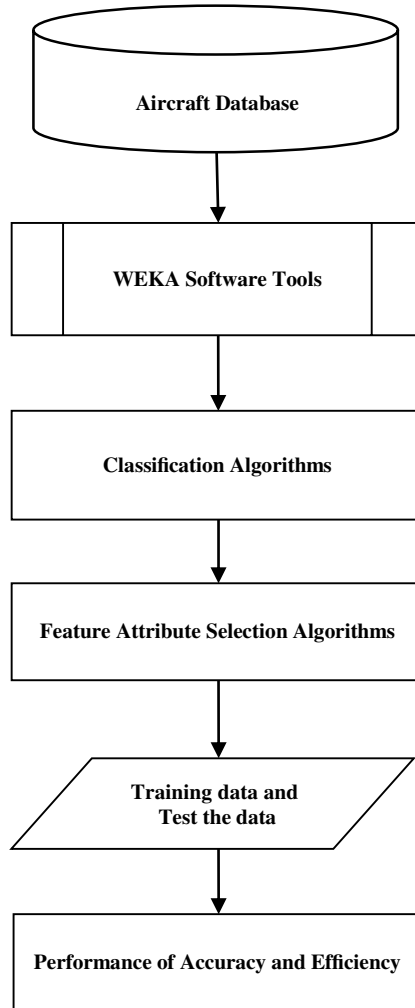
Figure 1. Block diagram of the proposed classification model.

Step 3:	Create compound classifier as the aggregation of particular classifiers.
Step 4:	Train the data using feature attribute selection algorithms to produce a classifier model.
Step 5:	Input test data set
Step 6:	Obtain test data and classify using the altogether built model.
Step 7:	Initiate query for Flight Delays using classifiers.
Step 8:	Analyze performance of accuracy and efficiency.
Step 9:	Output Result Analysis

Figure 2 describes the aircraft dataset along with pre-processing, which results in accurate performance and efficiency through classification algorithms and feature attribute selection methods. This model is proposed for improving different classification accuracy by combining the prediction of multiple classifiers. The different data mining classifiers like NB, Functions, Lazy, Meta, MISC, Rules and DT are used for classification.
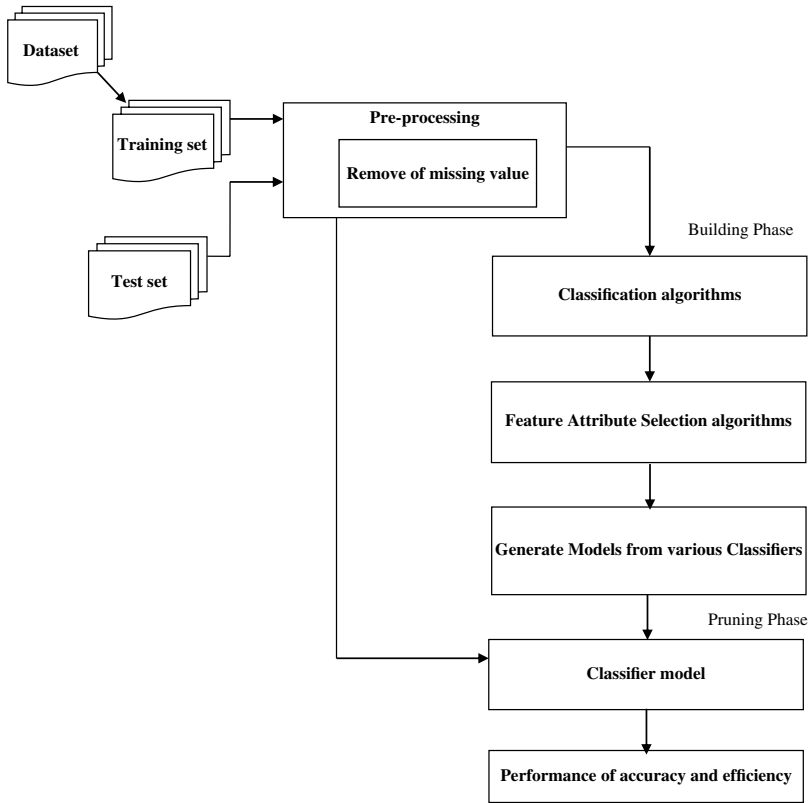
Figure 2. High level model of the system.

The classification models that are evaluated using the different performance metrics are Kappa statistic, Mean Absolute Error, Root Mean Squared Error, Relative Absolute Error and Root Relative Squared Error. A comparative study is carried out on the performances of the different classifiers, and after carrying out different feature selection techniques, the classifiers accuracy significantly improved.

$$\text{Kappa statistic} = \frac{\text{totalAccuray} - \text{randomAccuracy}}{1 - \text{randomAccuracy}}$$

Mean Absolute Error is a model evaluation metric used with regression models. The Mean Absolute Error of a model with respect to a test set is the mean of the absolute values of the individual prediction errors on overall instances in the test set. Each prediction error is the difference between the true value and the predicted value for the instance.

$$\text{Mean absolute error} = \frac{\sum \text{ni} = 1\text{abs}(\text{yi} - \lambda(\text{xi}))}{n}$$

Table 1
Confusion matrix for a two-class classifier

		Predicted	
		Negative	**Positive**
Actual	Negative	**a**	**b**
	Positive	**c**	**d**

where y_i is the true target value for test instance x_i, $\lambda(x_i)$ is the predicted target value for test instance x_i, and n is the number of test instances.

The performance of each feature selection method and classifiers model is evaluated by using statistical measures like accuracy, specificity and sensitivity. Receiver Operating Characteristics (ROC) parameters are used to compare the results of various classifiers. For the percentage of accuracy, sensitivity percentage and specificity percentage, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) expressions are used. The various formulas based on ROC parameters are given below:

Total accuracy is simply the sum of TP and TN, divided by the total number of items, that is:

$$\text{Accuracy} = \frac{(\text{True Positive} + \text{True Negative})}{(\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative})}$$

$$\text{randomAccuracy} = \frac{(\text{TN} + \text{FP}) * (\text{TN} + \text{FN}) + (\text{FN} + \text{TP}) * (\text{FP} + \text{TP})}{\text{Total} * \text{Total}}$$

$$\text{Sensitivity} = \frac{(\text{True Positive})}{(\text{True Positive} + \text{False Negative})}$$

$$\text{Specificity} = \frac{(\text{True Negative})}{(\text{True Negative} + \text{False Positive})}$$

A confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. Table 1 shows the confusion matrix for a two-class classifier. The entries in the confusion matrix have the following meaning in the context of our study:

- a is the number of **correct** predictions that an instance is **negative**,
- b is the number of **incorrect** predictions that an instance is **positive**,
- c is the number of **incorrect** of predictions that an instance **negative**, and
- d is the number of **correct** predictions that an instance is **positive**.

The *accuracy* (*AC*) is the proportion of the total number of predictions that were correct. It is determined using the equation:

$$AC = \frac{(a + d)}{(a + b + c + d)}$$

We proposed a different feature attribute selection evaluator method in this model. The number of features obtained in the dataset is very large. We have applied the number of different attribute evaluators CFS subset evaluator, consistency subset evaluator, gain ratio feature evaluator, information gain attribute evaluator, OneR attribute evaluator, principal components attribute transformer, relief attribute evaluator and Symmetrical Uncertainty Attribute Evaluator. The principal components attribute transformer is better than other attribute evaluators because of the accuracy is very high and minimum time execution.

4.0 DATA SOURCES AND DESCRPTION

For this study, we collected the data from the database of aircraft aviation. FAA officials, airlines, air traffic controller shortage, poor airport design, bad weather, heavy traffic and close proximity delays flights at several major airports in the eastern U.S. The application is done on number of datasets to compare the results. We are collecting data covering 1920 to 2015. In this manner, we used 5,112 data instances for various data mining algorithms. The data is given in a report layout with the following sections shown in Table 2. As is apparent from Table 2, component reports have a greater number of attributes. The aim of the analysis is to find the attributes that affect the warning levels of aircraft.

5.0 EXPERIMENTAL RESULTS AND DISCUSSION

We have applied the classification algorithms like that DT, NB, SVM, KNN and NN to the aircraft database dataset.[9] After the classified dataset, we used different feature selection techniques like that CFS, CS, GR, OneR, PCA, RF and SU to build classification models on the aircraft datasets with various selected subset of features. These classification models are evaluated in terms of the accuracy and efficiency performance metrics.

The major contribution is the study of the classification performance in terms of accuracy and efficiency with different feature selection methods. It is observed from the comparison of different classifiers based on the performance metric Kappa Statistic, Means Absolute Error, Root Mean Squared Error, Relative Absolute Error and Root Relative Squared Error on reduced features by different feature selection methods and also found the performance of metric results are reliable.

Table 3 shows the performance of accuracy with different Bayes classifiers such as Aode, Aodesr, BayesNet, HNB, NB and Waode on a large-scale aviation dataset. Table 4 shows the performance of efficiency with different Bayes classifiers such as Aode, Aodesr, BayesNet, HNB, NB, and Waode on a large-scale aviation dataset. Table 5 shows the performance of accuracy with different function classifiers such as Logistics, MLP, RBF and Simple Logistics. Table 6 shows the performance of efficiency with different function classifiers such as Logistics, MLP, RBF and Simple Logistics. Table 7 shows the performance of accuracy with different Lazy classifiers IB1, IBK and Kstar. Table 8 shows the performance of efficiency with different lazy classifiers IB1, IBK and Kstar.

Table 9 shows the performance of accuracy with different Meta classifiers: AdaBoost1, Attribute selected, Bagging, ClassificationviaClustering, ClassificationviaRegression, CVParameterSelection, Dagging, Decorate, END, FilteredClassifier, Grading, LogitBoost, MultiBoostAB, MultiClassClasifier and MultiScheme. Table 10 shows the performance

Table 2
Description of sample datasets used in application

Parameter name	Description
Abind	The challenging for the pilot to diagnose in flight
Aflalo	The side of the runway after landing long
Airatt	The aircraft will be put into a round out attitude shortly before it would otherwise contact the ground
Airbrot	The Burnt-out aircraft is clustered on the aft section of the flight deck, clear of the fire area
Arspd	The highest airspeed attained by an aircraft of a particular class
Altplm	The standard nominal altitude of an aircraft, in hundreds of feet
Apnopre	The pilot-interpreted make use of ground beacons and aircraft equipment such as VOR, NDB
Arhorfa	The view of heading indicator and artificial horizon after an in-flight vacuum failure
ATCerr	The service provided by ground-based controllers who direct aircraft on the ground
Autlad	The designed to make landing possible in visibility too poor to permit any form of visual landing
Autpidiso	The autopilot can control the aircraft while the pilot attends to other duties
Autpieng	The autopilot must be turned on using the Autopilot Engage Switch on the far left
Autthrot	The pilot to control the power setting of an aircraft's engines by specifying a desired flight characteristic
Bdwetr	The atmospheric conditions that comprise the state of the atmosphere in terms of temperature and wind
Belruele	The100ft above the touchdown zone elevation of the runway of intended landing and below that altitude
Brdstr	The bird strikes happen most often during take-off or landing, or during low altitude flight.
Bomrmnt	The key tactical control and administrative organisation for bombers in all theatres of operation.
Boucd	The pilot lost control of the aircraft landed hard and bounced on the runway during landing.
Brfa	The brakes failed.
Burfla	A light aircraft burst into flames after crashing into a building during an air show.
Cenln	The aircraft exactly on the centreline there will be bumping at each centre light.
Cirsf	Inertial reference systems (IRS) use ring laser gyros and accelerometers in order to calculate the aircraft position.
Climb	It is used to climb after takeoff until ready to leave the traffic circuit.

Table 2
Continued

Parameter name	Description
Corctrl	The aircraft structures are made of metal, and the most insidious form of damage to those structures is corrosion.
Crasd	To break in pieces violently; to dash together with noise and violence.
Crswnd	The component of wind that is blowing across the runway making landings and take-offs more difficult
Crstrfa	The structural cruise speed at which speeds must be below to avoid damage in turbulence.
Cvrinfo	It is used to record the audio environment in the flight deck of an aircraft for the purpose of investigation.
Damexpl	Explosions occurred below the aircraft impact levels in both towers prior to the collapses.
Dsnerr	Aircraft design is a compromise between many competing factors and constraints
Dirloctrl	The pilot to control the direction (left or right) of yaw about the airplane's vertical axis for minor adjustments.
Dsaprd	Many aircraft have gone missing without trace.
Dixmd	The displacement launched
Dme	To determine their distance from a land-based transponder by sending and receiving pulse pairs
Drmalfn	The accident was most likely caused by improper wiring and deficiencies in the door's design
Doendabld	The twin-engine aircraft must be able to fly after losing an engine.
Dwnwnd	The direction of the wind just like blown smoke
Durtkof	For light aircraft, usually full power is used during takeoff
Durtkofr	Usually the engines are run at full power during takeoff
Elfa	The electrical failures often result from interconnection breakdown between aircraft systems.
Emrland	The operation of the aircraft or involves sudden medical emergencies necessitating diversion to the nearest airport
Endruny	That end of the runway nearest to the direction from which the final approach is made.
Enfa	The engine failure is probably your worst fear as a pilot
Enmalfn	The engine failed on the way to town
Eqcolngsyplm	The electrical or electronic equipment on board an aircraft system problem.
Errinstrdng	The instruments in the cockpit of an aircraft that provide the pressure
Foaprd	To act as both an assertive individual and as a subordinate in a team atmosphere.
Fabldmgd	The fan blade damaged
Fagd	The progressive and localised structural damage that occurs when a material is subjected to cyclic loading.

Table 2
Continued

Parameter name	Description
Fctm	The Flight Crew Training Manual provides information and recommendations on manoeuvres and techniques.
Fdrinfo	The performance of various aircraft systems, as well as the aircraft's speed, altitude, heading and flight parameters.
Fr	The most serious in-flight emergency and must be brought under control as soon as possible.
Fstmdarcol	A mid-air collision is an aviation accident in which two or more aircraft come into contact during flight
Flplm	When used during takeoff, flaps trade runway distance for climb rate and reduce ground roll and the climb rate.
Fggr	The most frequent cause of surface visibility below 3 miles.
Fghvy	Heavy fog forced drivers to slow down
Fghl	This type of fog usually forms a good distance from the peak of the hill or mountain and covers a large area.
Fglgt	The fog that obscures less than 60% of the sky and does not extend to the base of any overhead clouds.
Fcdldng	The aircraft is forced to make a landing due to technical problems.
Frmruny	Defined rectangular area on a land aerodrome prepared for the landing and takeoff of aircraft".
Fulex	Exhaustion is when one quits running and falls into a heap, so too is fuel exhaustion with respect to an aircraft
Fulstar	There is a supply problem which either fully or partially prevents the fuel from reaching the engine.
Fultkexpd	Fuel tanks have been implicated in aviation disasters, being the cause of the accident or worsening it.
Fultkigd	Any process or event capable of causing a fire or explosion.
Fultklkd	The easy-to-fix to more difficult and expensive problems.

Table 3
Bayes Classifiers accuracy with whole aircraft dataset before applying feature selection methods

Classifiers	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy (%)
Bayes-Aode	3916	1196	76.60
Aodesr	**3939**	**1173**	**77.05**
BayesNet	3911	1201	76.50
HNB	3938	1174	77.03
Naive-Bayes	3911	1201	76.50
Waode	3925	1187	76.78

Table 4

Bayes Classifiers efficiency with whole aircraft dataset before applying feature selection methods

Classifiers	Kappa Statistic	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
Bayes-Aode	0.0281	0.3379	0.4104	93.91	96.77
Aodesr	0.0497	0.3366	0.4087	93.56	96.38
BayesNet	0.0253	0.3392	0.4117	94.28	97.07
HNB	0.0525	0.3346	0.4082	93.01	96.25
Naive-Bayes	0.0253	0.3395	0.4117	94.36	97.09
Waode	0.0342	0.3374	0.4102	93.77	96.72

Table 5

Functions Classifiers accuracy with whole aircraft dataset before applying feature selection methods

Classifiers	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy (%)
Logistics	3909	1203	76.46
MLP	3908	1204	76.44
RBF	**3913**	**1199**	**76.54**
Simple Logistics	3912	1200	76.52

Table 6

Functions Classifiers efficiency with whole aircraft dataset before applying feature selection methods

Classifiers	Kappa Statistic	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
Logistics	0.027	0.3387	0.4115	94.15	97.02
MLP	0.1098	0.3317	0.4146	92.20	97.77
RBF	0.0426	0.3366	0.4111	93.55	96.92
Simple Logistics	0.0223	0.3452	0.4145	95.96	97.75

Table 7

Lazy Classifiers accuracy with whole aircraft dataset before applying feature selection methods

Classifiers	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy (%)
IB1	3503	1609	68.52
IBK	**3917**	**1195**	**76.62**
Kstar	3915	1197	76.58

Table 8
Lazy Classifiers efficiency with whole aircraft dataset before applying feature selection methods

Classifiers	Kappa Statistic	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
IB1	0.0813	0.3147	0.561	87.49	98.89
IBK	0.0648	0.3333	0.4138	92.64	97.57
Kstar	0.0072	0.354	0.4155	98.41	97.97

Table 9
Meta Classifiers accuracy with whole aircraft dataset before applying feature selection methods

Classifiers	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy (%)
AdaBoostM1	3910	1202	76.48
AttributeSelected	3902	1210	76.33
Bagging	**3921**	**1191**	**76.70**
ClassficationviaClustering	3918	1294	74.68
ClassficationviaRegression	3910	1202	76.48
CVParameterSelection	3910	1202	76.48
Dagging	3909	1203	76.46
Decorate	3918	1194	76.64
END	3920	1192	76.68
FilteredClassifier	3920	1192	76.68
Grading	3910	1202	76.48
LogitBoost	3909	1203	76.46
MultiBoostAB	3910	1202	76.48
MultiClassClassifier	3895	1217	76.19
MultiScheme	3910	1202	76.48

of efficiency with different Meta classifiers: AdaBoost1, Attribute selected, Bagging, ClassificationviaClustering, ClassificationviaRegression, CVParameterSelection, Dagging, Decorate, END, Filtered Classifier, Grading, LogitBoost, MultiBoostAB, MultiClassClasifier and MultiScheme.

Table 11 shows the performance of accuracy with different MISC classifiers: Hyperpipes and VF1. Table 12 shows the performance of efficiency with different Misc classifiers: Hyperpipes and VF1. Table 13 shows the performance of accuracy with different Rules classifiers: ConjunctiveRule, DecisionTable, DTNB, JRip, NNge, OneR, Part, Prism, Ridor and ZeroR. Table 14 shows the performance of efficiency with different Rules classifiers Conjunctive Rule, Decision Table, DTNB, JRip, NNge, OneR, Part, Prism, Ridor and ZeroR.

Table 15 shows the performance of accuracy with different trees classifiers: ADTree, BFTree, DecisionStump, FT, Id3, J48, J48graft, LADTree, NBTree, RandomForest, RandomTree, REPTree, Simplecart and UserClassifier. Table 16 shows the performance

Table 10
Meta Classifiers efficiency with whole aircraft dataset before applying feature selection methods

Classifiers	Kappa Statistic	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
AdaBoostM1	0.0	0.3615	0.4204	100.47	99.13
AttributeSelected	0.0021	0.3594	0.4245	99.90	100.08
Bagging	0.0326	0.4126	0.4126	94.77	97.29
ClassficationviaClustering	0.0	0.2531	0.5031	70.36	118.63
ClassficationviaRegression	0.0	0.3597	0.4241	99.98	100
CVParameterSelection	0.0	0.3598	0.4241	100	100
Dagging	0.0	0.2397	0.4761	66.62	112.25
Decorate	0.0454	0.4438	0.4515	123.37	106.46
END	0.0255	0.3564	0.4228	99.05	99.69
FilteredClassifier	0.0255	0.3564	0.4228	99.05	99.69
Grading	0.0	0.2351	0.4849	65.35	114.34
LogitBoost	0.0169	0.3459	0.4153	96.14	97.93
MultiBoostAB	0.0	0.2396	0.4805	66.60	113.31
MultiClassClassifier	0.0132	0.3406	0.414	94.67	97.62
MultiScheme	0.0	0.3598	0.4241	100	100

Table 11
Misc Classifiers accuracy with whole aircraft dataset before applying feature selection methods

Classifiers	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy (%)
HyperPipes	**3910**	**1202**	**76.48**
VF1	2354	2758	46.04

of efficiency with different Trees classifiers ADTree, BFTree, DecisionStump, FT, Id3, J48, J48graft, LADTree, NBTree, RandomForest, RandomTree, REPTree, Simplecart and UserClassifier. Table 17 shows the performance of different classifiers after applying feature selection attribute evaluator methods.

Figure 3 displays the performance of accuracy for different classifiers with different feature selectors on a large-scale aviation dataset. Figure 4 displays the performance of accuracy for different classifiers on a large-scale aviation dataset. Figure 5 displays the performance of accuracy for different feature attribute selectors on a large-scale aviation dataset. Figure 6 displays performance of different classification models after applied feature selection methods with ROC chart. The DT classifier is better than other classifiers accuracy is very high and minimum time execution.

Table 12
Misc Classifiers efficiency with whole aircraft dataset before applying feature selection methods

Classifiers	Kappa Statistic	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
HyperPipes	0.0	0.4999	0.4999	138.96	117.88
VF1	0.0972	0.4926	0.4955	136.94	116.84

Table 13
Rules Classifiers accuracy with whole aircraft dataset before applying feature selection methods

Classifiers	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy (%)
ConjunctiveRule	3910	1202	76.48
DecisionTable	**3917**	**1195**	**76.62**
DTNB	3900	1212	76.29
JRip	3911	1201	76.50
NNge	3154	1958	61.69
OneR-B6	3905	1207	76.38
PART	3907	1205	76.42
Prism	3863	1185	75.56
Ridor	3908	1204	76.44
ZeroR	3910	1202	76.48

Table 14
Rules Classifiers efficiency with whole aircraft dataset before applying feature selection methods

Classifiers	Kappa Statistic	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
ConjunctiveRule	0.0	0.3598	0.4241	100	100
DecisionTable	0.0175	0.352	0.4194	97.84	98.89
DTNB	0.0407	0.3504	0.4175	97.39	98.45
JRip	0.003	0.3593	0.424	99.88	99.97
NNge	0.0569	0.383	0.6189	106.46	145.93
OneR-B6	0.005	0.2361	0.4859	65.63	114.57
PART	0.0498	0.3365	0.4136	93.52	97.52
Prism	0.0227	0.2347	0.4845	65.98	114.80
Ridor	0.0019	0.2355	0.4853	65.46	114.43
ZeroR	0.0	0.3598	0.4241	100	100

Table 15
Trees Classifiers accuracy with whole aircraft dataset before applying feature selection methods

Classifiers	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy (%)
ADTree-B10-E	3905	1207	76.38
BFTree	3911	1201	76.50
DecisionStump	3910	1202	76.48
FT	3908	1204	76.44
Id3	3919	1193	76.66
J48	**3920**	**1192**	**76.68**
J48graft	3921	1191	76.70
LADTree	3916	1196	76.60
NBTree	3909	1203	76.46
RandomForest	3913	1199	76.54
RandomTree	3916	1196	76.60
REPTree	3919	1193	76.66
SimpleCart	3917	1195	76.62
UserClassifier	3910	1202	76.48

Table 16
Trees Classifiers efficiency with whole aircraft dataset before applying feature selection methods

Classifiers	Kappa Statistic	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
ADTree-B10-E	0.0188	0.4043	0.4297	112.395	101.31
BFTree	0.027	0.3435	0.4169	95.47	98.31
DecisionStump	0.0	0.3526	0.4199	98.01	99.01
FT	0.0165	0.2896	0.4495	80.49	105.99
Id3	0.071	0.3316	0.4136	92.18	97.53
J48	0.0255	0.3564	0.4228	99.05	99.69
J48graft	0.025	0.3562	0.4227	99.01	99.67
LADTree	0.0413	0.347	0.4154	96.45	97.94
NBTree	0.0237	0.3413	0.4139	94.86	97.58
RandomForest	0.0829	0.3336	0.4129	92.73	97.36
RandomTree	0.0667	0.3327	0.4146	92.48	97.75
REPTree	0.0326	0.3413	0.414	94.87	97.61
SimpleCart	0.0319	0.3421	0.4149	95.10	97.83
UserClassifier	0.0	0.3597	0.4241	99.98	100

Table 17
Performances of different classifiers after applying feature selection attribute evaluator methods

	CFS	Chi-square	CS	GR	IR	OneR	PCA	RF	SU	SVM	AVG
NB	93.03	96.94	98.14	93.03	96.94	99.31	99.82	96.96	96.94	96.96	96.80
Functions	92.99	96.94	98.12	92.99	96.94	99.31	99.86	96.88	96.94	96.94	96.79
Lazy	93.03	96.96	98.16	93.03	96.96	99.31	99.86	96.96	96.96	96.94	96.81
Meta	93.03	96.96	98.16	93.03	96.96	99.31	99.86	96.96	96.96	96.96	96.81
Misc	93.03	96.96	98.16	93.03	96.96	99.31	99.86	96.96	96.96	96.96	96.81
Rules	93.03	96.96	98.16	93.03	96.96	99.31	99.86	96.96	96.96	96.96	96.81
Trees	93.04	96.96	98.16	93.03	96.97	99.32	**99.85**	96.96	96.96	96.97	**96.82**
AVG	93.02	96.95	98.15	93.02	96.95	99.31		96.94	96.95	96.95	
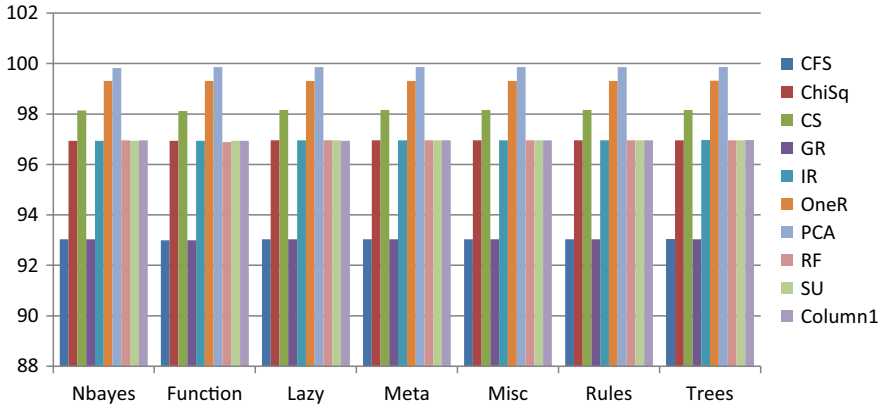
Figure 3. Performance of different feature attributes selectors with classifiers.
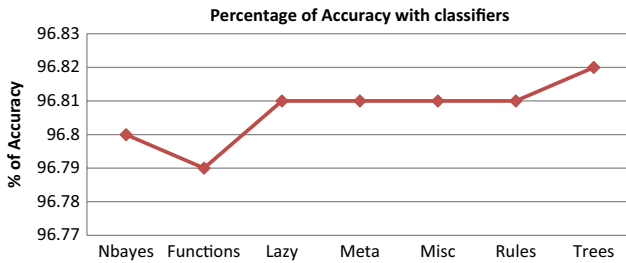

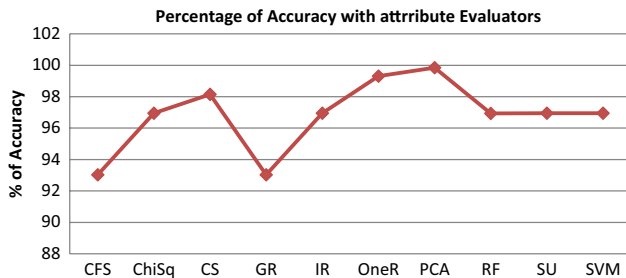Figure 4. Performance of different classifiers with accuracy.

Figure 5. Performance of different feature attributes evaluators with accuracy.

6.0 CONCULSION

In this study, different classifiers are used for the classification of aircraft data. We have studied the performance of different feature attribute selection methods. The main contribution of this study is to evaluate the performance of different classification algorithms like DT, NB, SVM, KNN and ANN on aviation aircraft data. This paper scrutinises the importance of feature attribute selection methods for improving the performance of different classification methods. It is found that difference performance accuracy and efficiency can be found with
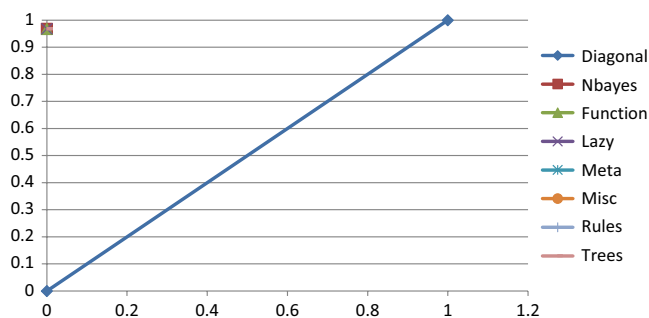
Figure 6. Performance of different classification models after applied feature selection methods with ROC chart.

the whole aircraft dataset before applying feature selection methods. After applying feature attribute selection methods, the WEKA tool increases the performance of different classifiers. We determined that the DT based on classifier and the PCA based on feature selector attribute evaluators are the best solution for aircraft analysis of flight delays using different data mining algorithms.

In this paper, we have explored different classification models and various evaluation attribute feature selection methods. By measuring the performance of the models using real data, we have seen interesting results on the predictability of the delays. The best delay prediction method appeared to be the most specific one, which considered all the combinations of categorical parameters and a condition on the arrival hour. The performances of the models were challenging to evaluate due to the variety of measures used and the different parameterisations adapted to them. However, the predictions obtained appeared to be better than the one seen in the literature. The classification algorithm was a very interesting method to learn and manipulate. Being a PCA method, it can be used each time we want to reconstruct a probabilistic model from some observations.

7.0 FUTURE WORK

The results of this qualitative research have significant implications for future research into the phenomenon of delay and for the design of policies aimed at solving this significant problem. First, these reports from the field highlight the fact that delay is as much a problem of unreliability as it is of longer trip times. When the air transportation system is plagued by delay, travellers become less able to predict when they will arrive, and so they become less able to plan their trips efficiently.

We could also model the phenomenon more precisely if instead looking at more than just the distribution of past data. We can, for instance, build separate models per time period, per type of aircraft, per airline and per region, and then group them into a general model. This way, we might be able to predict the delays of a new flight without needing several months of data to build a prediction model. Another step forward would be to generalise the model to flights of the entire world, or at least to exploit more data sources, to build more complete predictions. Finally, the most interesting step would be to integrate such a model into a flight booking tool, to provide the delay prediction to future passengers, even if this would require a strong confidence in the information provided, considering the possible impact in terms of reservations.

ACKNOWLEDGEMENTS

REFERENCES

1. WU, C. Inherent delays and operational reliability of airline schedules. *Journal of Air Transport Management*, 2005, **11**, (4), pp 273–282.
2. ERIC, R. and CHATTERJI, B. Analysis of aircraft arrival and departure delay characteristics. NASA Ames Research Center, Moffett Field, CA 94035-1000, Integration, and Operations (ATIO) Technical, Los Angeles, California, 2002, pp 1–3.
3. AISLING, J. and REYNOLDS-FEIGHAN, A.J. An Assessment of the Capacity and Congestion Levels at European Airports. 39th Congress of European Regional Science Association, University College Dublin, Ireland. 24–27 August 1999.
4. EUROCONTROL. *Delays to Air Transport in Europe*, 2001, Annual Report.
5. PHILLIPS, D., and STEELE, P. A Graphical Tool to Visualize Predicted Minimum Delay Flights. *Journal of Computational and Graphical Statistics*, 2009, **20**, (2), pp 294–297.
6. WANG, Y., KLEIN, A., and JEHLEN, R. Eighth USA/Europe Air Traffic Management Research and Development Seminar (ATM) Modeling Flight Delays and Cancellations at the National, Regional and Airport Levels in the United States Banavar Sridhar, 2009.
7. SMITH, A. and SHERRY, L. *Decision Support Tool for Predicting Aircraft Arrival Rates, Ground Delay Programs, and Airport Delays from Weather Forecasts*. Center for Air Transportation, Systems Research George Mason University, Fairfax, VA, 2008, pp 1–4.
8. CHRISTOPHER, A.B.A. and BALAMURUGAN, S.A. Prediction of warning level in aircraft accidents using data mining techniques. *The Aeronautical Journal*, 2014, **118**, (1206), pp 935–952.
9. Aircraft accident database dataset retrieved from http://www.planecrashinfo.com/database.htm.