



## Journal of Internet Banking and Commerce

*An open access Internet journal (<http://www.arraydev.com/commerce/jibc/>)*

*Journal of Internet Banking and Commerce, December 2012, vol. 17, no.3  
(<http://www.arraydev.com/commerce/jibc/>)*

### Validation and Classification of Web Services using Equalization Validation Classification

---

**ALAMELU MUTHUKRISHNAN**

**Assistant Professor(Sr.G),Information Technology Department, B.S.Abdur Rahman University.**

**Postal Address: Assistant professor(Sr.G), Information Technology Department, B.S.Abdur Rahman University, Seethakathi Estate, GST Road, Vandalur, Chennai-48,India.**

**Organizational Website: [www.bsauniv.ac.in](http://www.bsauniv.ac.in)**

**Email: [M.Alamelum@gmail.com](mailto:M.Alamelum@gmail.com)**

Ms. M.Alamelu is working as a Assistant Professor(Sr.G) in the Department of Information Technology, B.S.Abdur Rahman University, Chennai.She completed her undergraduate Degree B.E(CSE) from St.Xavier's Catholic College of Engineering, Nagercoil in the year of 2002 and her post graduate degree M.E(SE) in Sri.Ramakrishna Engineering college, Coimbatore in the year of 2005.Currently she is doing her research in the area of Service Oriented Architecture at Anna University of Technology, Coimbatore. Her research interest includes Web services, Artificial Intelligence and Neural networks.

**AM JAFFER MOHAMED ZUBAIR RAHMAN**

**Professor, Computer Science and Engineering Department, Al-Ameen Engineering College.**

**Postal Address: Principal, Professor, Computer Science and Engineering Department, Al-Ammen Engineering college, Karundevan palayam,Nanjai Uthukili, Erode- 638 104, Tamil Nadu, India.**

**Organizational web site: [www.alameen.ac.in](http://www.alameen.ac.in)**

**Email: [mdzubairrahman@gmail.com](mailto:mdzubairrahman@gmail.com)**

Dr. A.M.J.Md Zubair Rahman working as a principal at Al-Ameen Engineering College Erode. He has completed his M.S.,(Software Systems) in BITS Pilani, India in 1995.Then he completed M.E.,(Computer Science & Engineering) in Bharathiar University, Tamilnadu,India in 2002.

He obtained his Ph.D in the field of Data mining from the department of Computer Science Engineering, Anna University, Chennai in the year 2010. He has 20 years of teaching experience. He has published 15 articles in International / National journals.

---

## **Abstract**

In the business process world, web services present a managed and middleware to connect huge number of services. Web service transaction is a mechanism to compose services with their desired quality parameters. If enormous transactions occur, the provider could not acquire the accurate data at the correct time. So it is necessary to reduce the overburden of web service transactions. In order to reduce the excess of transactions from customers to providers, this paper propose a new method called Equalization Validation Classification. This method introduces a new weight-reducing algorithm called Efficient Trim Down algorithm to reduce the overburden of the incoming client requests. When this proposed algorithm is compared with Decision tree algorithms of (J48, Random Tree, Random Forest, AD Tree) it produces a better accuracy and Validation than the existing algorithms. The proposed trimming method was analyzed with the Decision tree algorithms and the results implementation shows that the ETD algorithm provides better performance in terms of improved accuracy with Effective Validation. Therefore, the proposed method provides a good gateway to reduce the overburden of the client requests in web services. Moreover analyzing the requests arrived from a vast number of clients and preventing the illegitimate requests save the service provider time.

**Keywords: Equalization Validation Classification (EVC), Efficient Trim Down (ETD), Combined Group Classifier (CGP), Request Recognizer (RR)**

© Alamelu Muthukrishnan and AM Jaffer Mohamed Zubair Rahman, 2012

---

## **INTRODUCTION**

Normally transactions involve the communication between one or more number of resources. Based upon the keenness of the resources, the resources are locked or unlocked by the systems. Web services transactions, play a vital role in day to day life. For example, Ticket reservation, Hospitality, E-business, and Organizations mostly depends upon the online web service transactions. A complex business process involves multiple web service compositions to utilize the dependable resources. Such utilization of those resources involves multiple transactions from one set of input services to the other set of output services. In way of transactions, the data transferred from one source to the other source to be confidential, consistent and should be reliable. Consider a travel booking web service in which ticket booking, Hotel reservation and vehicle reservation are involved. Here one set of services are given as input to other services. If any one of the services gets affected, then it will reflect the whole services.

No more data will be transmitted from one service to the other service. So transaction plays an important role in web services.

Various challenges have been discussed in the past for web service transactions. Most of them are based on Time, Reliability, Trust, Risks and Isolation. The first issue starts with the time constraint in web service transactions. In business point of view, web services transactions are more complex with time. Respect to the model based Atomicity, Consistency, Isolation and Durability (ACID) properties is a standard model generally used for the business transactions. Web service transactions, are loosely coupled and the transactions may extend over hours or a day. Due to the extended period of time, transactions may get locked and will not contribute fully to the business processes during such periods.

The second issue is related with the reliability of transactions as desired by the transaction coordinator and resource manager. Since the loosely coupled systems are synchronous, connection oriented protocols are suitable maintaining the communication between the transaction coordinator and resource manager. If any communication failure occurs, the resource manager or the transaction coordinator can no longer access the service. Comparing with the tightly coupled transactions the loosely coupled transaction messages must be more reliable to reduce the faults.

The third issue to be discussed is the trust that ensures that the resource to be more credible to the accessed services. Therefore the resource manager should provide surety to the resources. The best solution to handle these issues is to define a timeout period with the resource manager to make use of the service with a given period of time.

The fourth issue is the risks expected during the web service transactions. Most risks are involved during the transactions of web service. For example, in an airline reservation service, a person books the airplane ticket before two or three months. Due to the circumstances, the person may not be able to travel and hence the amount he/she paid is not so refundable. The criteria here is the resource has to detect and prevent the misdeed happens in such situations. Under such transactions are made, the web service has to define a restricted policies to avoid such risks. The last issue confers with the Isolation of transactions. As for the overflowing of requests, the requests should be flowed through a single queue of processing. If the Isolation happens without locking process will reduce the overflowing of requests during transactions.

Thus over considering the above issues the web service transactions has to be more reliable and timeliness. In this paper proposing an Equalization Validation Classification (EVC) method to solve the issue of overburden of incoming requests from multiple clients. The provider can lighten up the overburden of the incoming requests by using this Equalizing Validation Classification method. In this method, an Efficient Trim Down (ETD) algorithm analyzes the data with the measures of empirical estimation, complexity and input parameters. Moreover, it is matched with the Decision tree algorithms like (J48, Random Tree, Random Forest, AD Tree) and is found that this proposed algorithm produces better accuracy than the existing algorithms.

## LITERATURE SURVEY

The existing works define the different new technologies and approaches for web services classifications. The literature reviews here discuss the relevant fields of web service classification and the various methodologies and approaches.

Stephen S. Yau, Fellow (2008), introduces a Secured sharing repository for various shared services. This privacy-preserving repository has the specialized feature than other central repositories since they have designed it based on the user integration requirements with central control and precise results. They carried out the Evaluation of the framework with the Query framework named Query plan Executer with Query plan wrapper. So as with user-integrated requirements, this repository can easily decompose and discover the data's from existing services and finally produce to the users. The main advantage of their repository is that it uses additional aggregate functions and encryption schemes to prevent shared data accessing.

Sattar Hashemi, et al (2009), focused on the sophisticated advantages of One-Versus-all (OVA) classifiers to classify the streaming data. Some of the highlighted issues discussed in their paper are low error correlation, the adopting of new class labels and foreword of new OVA scheme to reduce the imbalanced class distributions. OVA is the new field of classification and produce a faster accuracy and updating than other classifiers.

Qianxiang Wang, et al (2009), introduced an approach called online monitoring for web service requirements in which the foremost requirements are collected from the users. They are user request, resource, response, domain application and management. The classifications of the user requirements are collected with the two components like monitoring code deployment manager and monitoring code generator. Also the monitoring process was connected with the agents to collect data that responsive to the requirements.

Dimitrios Skoutas et al (2010), introduced a method for ranking and clustering of the web services based on the web service search results. For ranking, the authors implemented three different algorithms on the search results and two different algorithms to select the agent services. With this method, the searching results efficiency was improved quickly than the existing algorithms.

Michael von Riegen, et al (2010), delivered a framework called TrackG to improve the autonomous coordination in a distributed systems. This paper explains how to handle the drawbacks of WS-Business Activity and WS-Coordination. The Predicate rules are assigned to participants and complicated process. A rule engine was used by the authors to evaluate these rules where the coordinators can control the process only in an autonomous way. With this new technology, rules are useful to prevent the difficulties in web service transactions.

Claudio A. Ardagna, et al (2011), provide the XACML (eXtensible Access Control Markup Language) base access control framework which supports access control mechanism for open web based systems. The main discussions of their paper are XACML, Credential, Abstraction, Recursion and Support of Dialog.

Jung-Yi Jiang, et al (2011), proposed a fuzzy similarity self (FFC) base clustering algorithm for text classification. Where the similar set of data is grouped into a set of clusters. Each cluster is characterized by mean and standard deviation. The weighted combinations of the words also calculated in a cluster. Their algorithm adapted by many researchers in the field of distributed word clustering. Moreover, it is applicable to the relevant fields like image processing, web mining and data sampling.

In summary, the existing methodologies discuss the general issues of web service transactions with the certain limitations. With the analysis of existing systems, the proposed method introduces a new algorithm called Efficient Trim Down (ETD) to reduce the overburden of the incoming requests. If the vast number of clients invokes service providers, it process gets delayed due to the overloading of incoming of data. The incoming data accuracy is measured with the Empirical estimation and Complexity.

### EQUALIZATION VALIDATION CLASSIFICATION METHODOLOGY

The proposed classification method, the Equalization processing is performed with a set of processing steps. Figure.1. shows the Equalization.

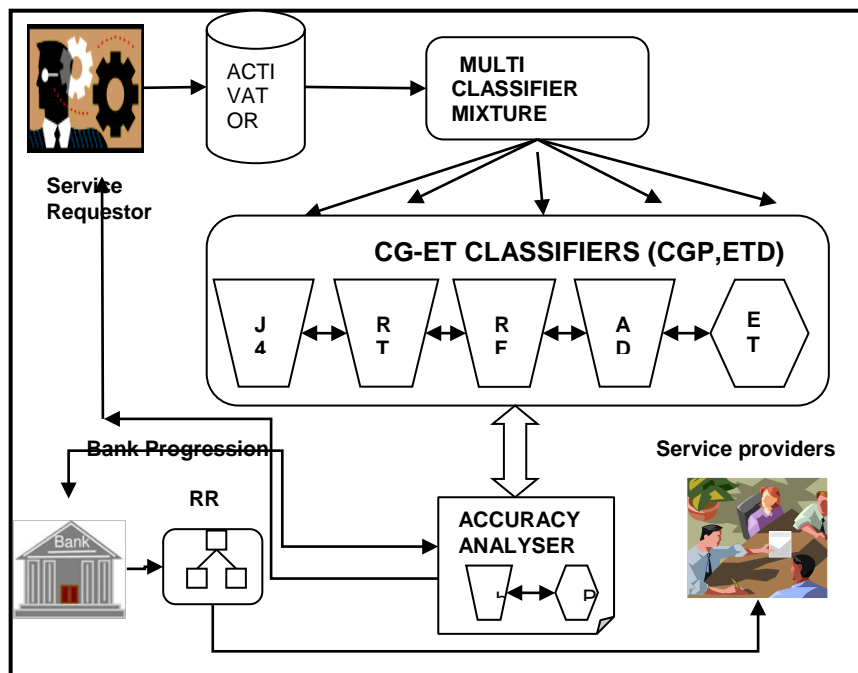


Figure.1. Equalization Validation Classification (EVC) method

The EVC method has listed out the set of different processing layers

1. **Service Requestor.**
2. **Activator.**
3. **Multi Classifier Mixture (MCM)**
4. **CG-ET Classifiers (CGP,ETD): { C1-J48, C2-Random Tree, C3-Random Forest,C4-ADTree,C5-Efficient Trim Down (ETD) }**

5. **Accuracy Analyzer**
6. **Bank progression**
  - 6.1 **Request Recognizer (RR)**
7. **Service Providers**

The process starts from the **Service requestor**. The requestor is a program or service that invoked from the clients. The service requestor may either act as a service provider. To purchase a product through online, web services takes an important role with service requestors or clients. A service requestor has the ability to identify the appropriate service provider based on their service offers, quality and cost. The providers offer various attractive offers to the customers; if the customers satisfy the stated offers they prefer the services. The proposed method discusses about how the web service transactions saturated with the overflow of requests. The process is initiated with the customer purchasing. If the vast number of transactions occurs at a time the processor called Activator collects it. When a different set of customers invoke a suitable provider, the authenticated details are collected from the activator. The main role of this activator is to collect the valuable credential information's from the multiple customers. The activator has the details of customer such as shop code, bill no, cardno and amount. If these valuable information's are generated, then only the Activator distributes to the **MCM mixture**. The collected Group classifiers define the Interface connectivity to the web service clients. The set of connection is interconnected in this mixture only. After the interface connectivity the incoming data are passed to the **CG-ET (CGP-ETD) Classifiers**. The CG are defined as **CGP (Collective Group of Classifiers)** is the category of (**C1-J48, C2-Random Tree, C3-Random Forest, C4-ADTree**). The CGP classifiers comparative results are matched with the ETD (Efficient Trim Down) classifier algorithm.

The CG-ET classification accuracy results are compared and passed to the **Accuracy Analyzer**. A bulk transitioned data are analyzed and the accuracy of each classifier algorithm is analysed. The accuracy analyzer clarifies that the ETD algorithm classification has the better accuracy and efficient validation than the existing algorithms. The analyzed reports state that the proposed algorithm provides better accuracy rate produced than the previous algorithm that is nearly equal to 100%. After validation, the **AA** passes the accuracy measure of credential information's to the **Bank Progression**. Either the Bank is a nationalized or internationalized, it must be tied up with the diverse service providers. When a user wants to access the particular provider, the user gets registered with the provider with the minimum bank balance amount. With the authorized registration, the client details are automatically transferred to the bank process. When the required accuracy is not met from the Accuracy Analyzer, the bank process will send an "**Invalid request**" message to the customers/clients. The matched and evaluated data are finally transmitted to the **Service providers**. As with the user initial registration, the accurate and validated data are transferred to the providers.

## **PROPOSED METHOD (EVC) PROCESSING**

### *Activator*

Activator is the starter-collector to collect the bulk of data in web service transactions. The collected data have the credential information like.. (shop code, bill no, cardno, amount). Before providing these information, the user should have the sufficient account balance to purchase the products. Such information's are collected in the activation

process and passed to MCM mixture.

### *Multi Classifier Mixture (MCM)*

It is an Interface Connectivity to the Classifiers and the service clients. The credential data are flowed from the activator to the mixture. Such as, MCM is to make an interface from the Activator to the set of classifiers like (GetJ48, Random Tree, Random Forest, AD Tree) and the proposed classifier of ETD (Efficient Trim Down). The efficient parameters are defined in this interface. The processing of MCM mixture is shown below:

### **MCM processing**

- Step 1:** Check for satisfaction of Requirements.
- Step 2:** If the authenticated data (`objweka.PreAuthenticate = true`) is true then the Credential data get passed to the classifiers.
- Step 3:** Service the clients to purchase items, with the credential informations of shopcode, Billno, cardno and amount. Here the, credential information results are Concatenated with the classifiers using Weka classifier using the following code.
- ```
CentralizedBank.weka.WekaClassifierService objweka = new
CentralizedBank.weka.WekaClassifierService();
objweka.PreAuthenticate = true;
objweka.Credentials= System.Net.CredentialCache.
DefaultCredentials;
```
- Step 4:** The bulk of credential data is (`objweka.getAllResult1(shopcode,billno,cardno,Amount)`) reterived from the customers and passed to the classifiers of CGP and ETD.

### *CG-ET Classifiers [ CGP(C1,C2,C3,C4), ETD ]*

After the collection of variables from the **MCM**, the processed data are next forwarded to the differencing Classifiers of **((CGP)->(C1,C2,C3,C4),ETD)**, Where the **C1** denotes J48, **C2** denotes Random Tree, **C3** denotes Random Forest and finally **C4** denotes AD Tree. The term ET denotes the proposed classifier of **ETD (Efficient Trim Down)** Classifier. In this CG-ET Classification, the vast amounts of data are accurately classified with these classification algorithms. The bulk credential data such as shopecode, billno, Cardno and amount are passed to CGP classifiers and also to the ETD classifier. Each classifier have the different parameters. The J48 classifier takes the bulk training data sets, bulk credential data's are augmented with the vector nodes. From this, the highest normalized base is identified. Once the base nodes are identified, the sub nodes are added to the base nodes.

In the Random Tree classifier, the classification the multiple decision trees are formed randomly. The variables (shopecode, billno, Cardno,amount) are assigned to the tree nodes in a random way and the filtering of the nodes are recorded during tree class distributions. In Random Forest classification, the Input variables are compared with the classifier variables. After the comparison, the input variables (shopecode, billno, Cardno, amount) are assigned randomly. In the Random Forest classification pruning is not applicable for the exceed limit. The AD Tree has the major of two nodes called Decision and precision nodes.

Decision node present as a parent node the variables (shopecode, billno,

Cardno,amount) are assigned in this node only. And the precision node as the sub node has the Instance values from the decision nodes.

The four Decision tree algorithms are evaluated with bulk credential information. The classification accuracy and validation results obtained are then matched with the proposed **ETD** classifier. The following decision algorithm shows the credential information dispensation with the CGP classifiers.

### **CGP Classifiers (CGP-> (C1, C2, C3, C4)) algorithms**

#### **CGP1: (C1) J48 Classifier:**

1. Read the training Data.
2. Augment the vector nodes with the dataset of shope code, bill no, Card no, Amount.
3. Identify the highest normalized base.
4. Base recursions are defined and add the sub nodes to the base using the following code.

```
String ret= WriteFile(shopecode,billno,CardNumber,amount);  
String[] args={"CLASSIFIER","weka.classifiers.trees.J48","U",  
"FILTER","weka.filters.unsupervised.instance.Randomize","DATAS  
ET","iris.arff"
```

#### **CGP2: (C2)RandomTree:**

1. Construct multiple Decision Trees.
2. Randomly allocate the parameters of shopecode,billno,Card no, amount into the tree.
3. Filtered nodes are identified and tree class distributions are recorded with the following code.

```
String ret= WriteFile(shopecode,billno,CardNumber,amount);  
String[] args={"CLASSIFIER","weka.classifiers.trees.RandomTree",  
"FILTER","weka.filters.unsupervised.instance.Randomize","DATASET",  
"iris.arff"};
```

#### **CGP3: (C3) Random Forest:**

1. Read the training data set and define the classifier variables.
2. Check the Input variables shope code, billno,Cardno, amount with the classifier variable.
3. Input variables should less than the classifier variables.
4. Random allocation provide for the Input variables.
5. Pruning is not permissible. And the following code describes the classification.

```
String ret= WriteFile(shopecode,billno,CardNumber,amount);  
String[] args={"CLASSIFIER","weka.classifiers.trees.RandomForest",  
"FILTER","weka.filters.unsupervised.instance.Randomize","DATASET","iris.arff"};
```

#### **CGP4: (C4) AD Tree:**

1. Define the Decision nodes and precision nodes.
2. Assign the Inputs shopecode, billno, cardno, amount to decision nodes.
3. Calculate the Instance based on the traverse of the prediction nodes. Code depicts the processing of classification.

```
String ret= WriteFile(shopecode,billno,CardNumber,amount);
```



```
String[]args={"CLASSIFIER","weka.classifiers.trees.ADTree",
"FILTER","weka.filters.unsupervised.instance.Randomize","DATASET","iris.arff"};
```

#### *Efficient Trim Down (ETD) Classifier*

Parallel to the evaluation with the CGP classifiers, the ETD classifiers are also processed. The input data was evaluated based on the empirical estimation, complexity and the error rate. Based on this algorithm, the Efficient Accuracy (EA) of the bulk data was calculated as per one of the risk factors in the data. For example, the previously analyzed sample transaction data are taken and matched with the bulk input data. From that the predicted results the risk factors are identified for the input data. The second process starts with the complexity analysis of the input data. The complexity of the bulk data is calculated with the number of records from the clients and the accuracy measure of the collected data. ie ..  $h(\log 2m^*a/h+1) - \log(n/4)$ . Where m is number of records input from the client requests.

The measure of complexity and empirical estimation of the vast data are estimated based on the number of records. How much of count given is divided with the factor of  $(m^*a) ^ (1/2)$ ? If more than the unlimited count is calculated then the processes of scheming is stopped. So, that the accurate measure of the data is presented and the unlimited transactions are also reduced with this classification algorithm. To improve the accuracy of the incoming data, EFD algorithm operates with the formula

$$\text{Efficient Accuracy (EA)} = R_{\text{emp}}() + (h(\log(2m^*a/h + 1) - \log(n/4)) / (m^*a))^{(1/2)}$$

#### where

$R_{\text{emp}}$  - Is the Empirical estimation of Risk factors

$m$  - Number of records Inputted from the client Requests

$h$ - Complexity of classification

**logarithm(log)** - Used to define the accuracy of the collected data's. For eg in probability distribution a value 2 have 1.414 and 3 have 1.752.

$a$ -is the stated Error rate of Input attributes , Where the number of attributes is directly proportional to the error risks. When the number of attributes increases, the error rate and the accuracy are also increased.

The following Rules demonstrate the procedural steps of the ETD Classifier algorithm.

#### **ETD (Efficient Trim Down) Classification Algorithm**

- R1: Define the Inputs shoppecode, billno, CardNumber, amount with the value of n.**
- R2: Calculate the Empirical estimation based on the risk factors.**
- R3: Match the predicted data with the Input data.**
- R4: Estimation of risk factors defined with Ramp( ) ->{ Emperical Estimation for risk factors}.**

- R5: Complication of the data is analyzed with the complexity of “h”.**  
**R6: Identify the Error rate and match the Input data with the measured probabilities.**  
It is defined as  $(\log(2m \cdot a/h + 1))$   
Where  $a \rightarrow$  is the stated Error rate of Input attributes which is directly proportional to the Error risks.  
**R7: Estimate the probability estimation with respect to the Input variables. and assume the probability ratio to produce accurate error rate of  $\log(n/4)$ .**  
**R8: Finalize the total count of Input records with  $m/a$ .**  
**R9: Calculate the logarithm probability with parameter of  $\frac{1}{2}$ .**  
**R10: Finally, find the better accuracy using Efficient Accuracy (EA).**

#### *Accuracy Analyzer (AA)*

The Evaluated classified data from the **CG-ET** are validated and analyzed with the Accuracy Analyzer. Such as the input data from the MCM is classified and forwarded to the CG-ET Classifiers. At first, the set of data which get into the MCM with the collective values of (billno, shope code, password and amount). Then these credential information are then fed into the **CG-ET**. Based on the different classifiers, the data are input and validated with the classifiers namely J48, Random Tree, Random Forest, AD Tree algorithms and parallel to ETD (Efficient Trim Down) classifier. The best defined accurate classifications of data are then matched with the proposed EFD accuracy rate. Comparing with the existing algorithms, the new proposed algorithm ETD produces better accuracy with respect to Riskless empirical estimation, Complexity and the valid Inputs. The resultant comparative analysis data are shown in Tables.1 and Table.2.

#### *Bank Progression*

Bank progression is an important dispensation to make the money transactions to the service providers. The Service provider services are tie-up transactions with the particular bank which is either nationalized or Internationalized. If a customer wants to buy any products from the service provider, the customer must have a sufficient balance during the transaction. If the minimum balance is not defined, the transaction cannot be established for the customer purchasing. The Bank progression has a sub evaluation part called **Request Recognizer (RR)** to produce the split data in a bulker that reduces the member of slow transactions.

#### *Request Recognizer (RR)*

The Request Recognizer (**RR**) is the one which gets response from the bank. The data are validated and analyzed by the **AA** and further passed to the bank progression. The RR only gets the accurate data after the comparative evaluation from the **CG-ET** Classifiers. If the accuracy is not so met from the **AA** an Invalid response is intimated to the client. If, so the accuracy is met the bulk data get split in the RR. The vast confidential data are placed in a different bulker to make the transaction fast. With the **Request Recognizer** processing the bank transaction get trouble-free passing data to the service providers. Due to the formation of this recognizer, the over burden has been reduced.

The step of the proposed RR is shown below.

### RR Processing

- Step1:** Stream the accuracy and validation results from CG-ET Classifiers.
- Step2:** Compare the results in the Accuracy Analyzer (AA).
- Step3:** Pass the finalized accuracy results to the sub evaluation part - RR (Request Recognizer).
- Step4:** Bulk data get split in the RR with different bulker.
- Step5:** If Accuracy is not matched, invalid reply is sent from the bank processing.
- Step6:** Otherwise, finalized data pass to the customer in a Queue way. The processing is describe with the code.

```

if(obj.CommandQuery("Insert into
Tbl_TransactionDetails(AccNo,Amount,Date,ChequeNo) values(" +
Accountno + "," + Amount + "," + DateTime.Now.ToShortDateString()
+ "," + ChequeNo + ")"))
{
obj.CommandQuery("Insertinto Tbl_DepositDetails(AccNo,Date,Amount)
values(" + dsOrgFrom.Tables[0].Rows[0][0].ToString() + "," +
DateTime.Now.ToShortDateString() + "," + TxtAmount.Text.Trim() + ")");
obj.CommandQuery("Update Tbl_CheckDetails set Status=1 where
CheckNo=" +
ChequeNo + " and AccountNo_FK=" + Accountno + ");
Result = "Amount CreditSucessfully";
}
else
{
Result = "Error Occured";}}
else { Result = "Sorry Insufficient Balance";}}}}
else { Result = "Cheque Bounces";}}
else{
Result = "Invalid Cheque Account Number"; }}
else{ Result = "Invalid Credit Account No";}

```

### **RESULTS AND DISCUSSIONS**

The following result shows the experimental value ranges of the 200 bulk data for web service transactions. The vast number of users had their transactions with a particular service provider or multiple providers. Due to the over burden of transactions, most of the times, the provider could not establish the correct response to the clients. To reduce the over burden of transactions, this performance experimental results show the web service transactions with the bulk amount of data.

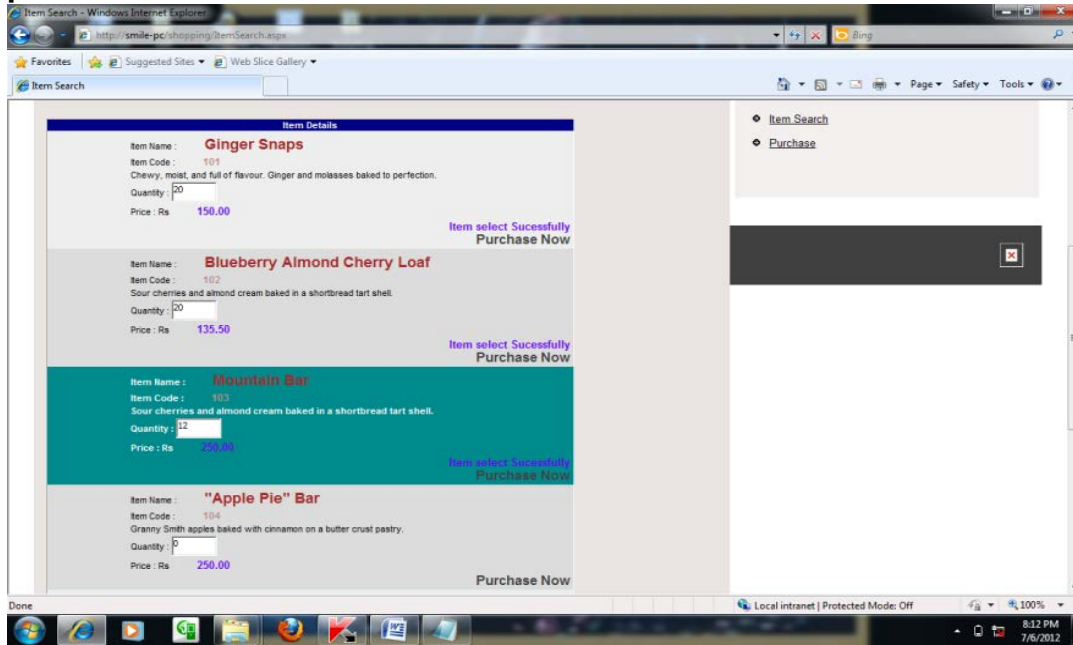
To experiment the web service transaction, a web service provider is created and published to access for the number clients. Here, the service provider publishes a service named REIN Departmental purchasing items to the customers.

The customers can buy their products with the REIN provider. The transactions have been tested with the 200 numbers of customers who access the service provider

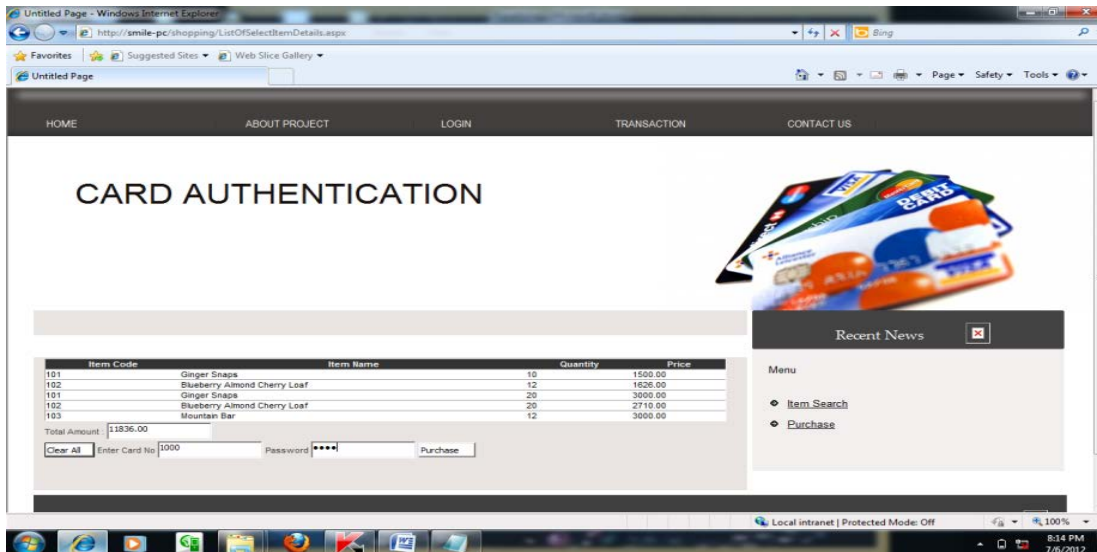
service. Initially, a bulk service customer's data is designed in a Web service Visual studio 2008 and their data are stored with a SQL database.

The below screen shots (1, 2, 3) display the web service transactions for purchasing, bank transaction and SQL data storage.

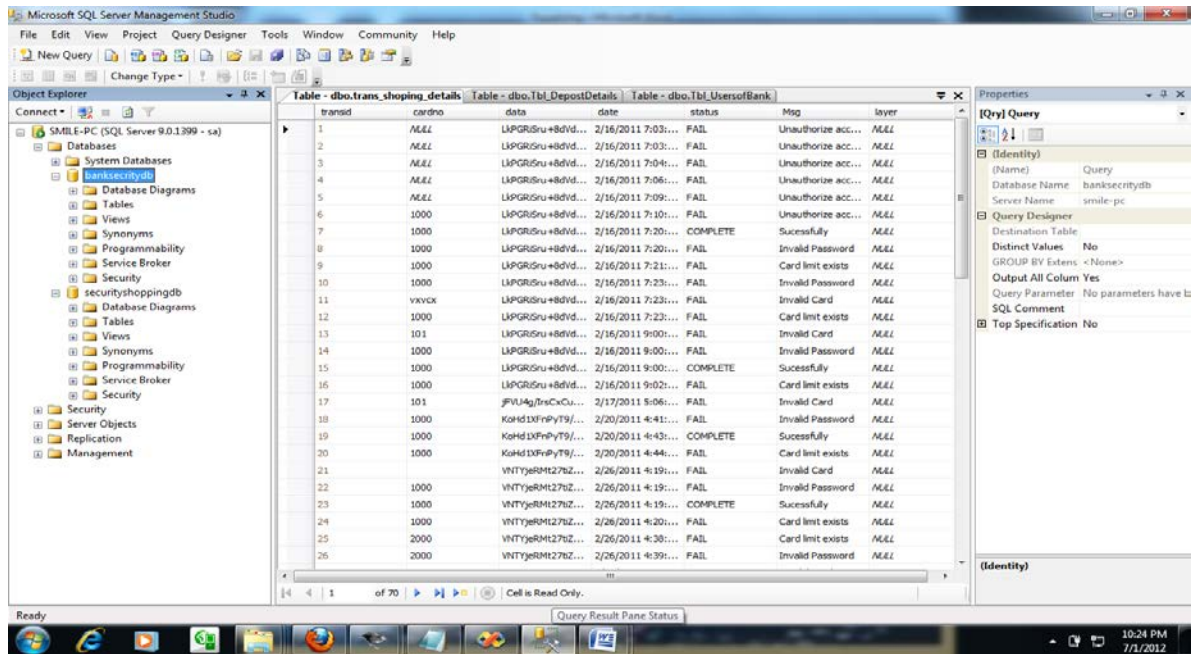
### Screen shot 1: Purchasing of items by service client from the service provider



### Screen shot 2: Card Authentication process in Bank Transaction



### Screen shot 3: Display of the web service client details from the SQL data base.



If any customer wants to buy a product through the REIN service provider, he/she must have the requirements satisfaction with the REIN provider. Also the REIN provider has all its customer related transactions with a bank known as YBN centralized bank.

It is assumed that the bank has the tie-up with the service provider which will define all transactions to the service provider. If the classified data from the classifier has not met the accuracy, an immediate response is sent to the service customers. In this process the REIN provider's time is saved and unauthorized credential data are not evaluated.

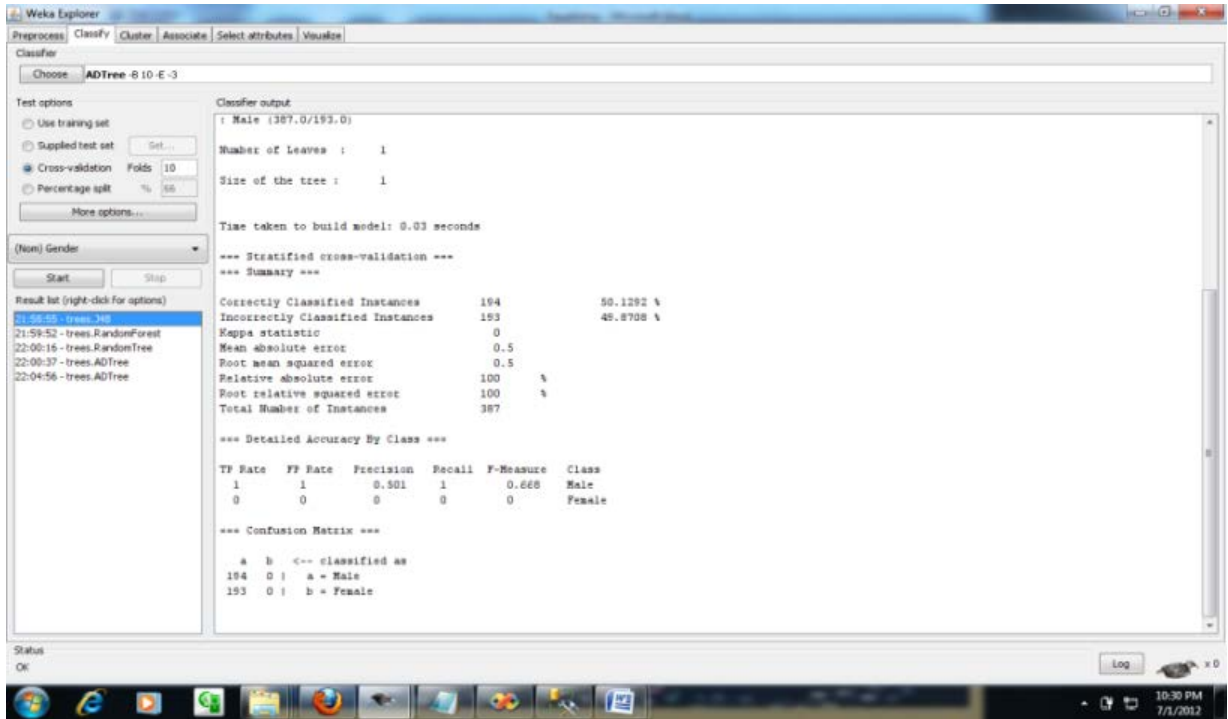
*Dataset Classification*

As with the creation of the web services, the accuracy and response of the bulk data are classified with the Classifiers. Here the decision tree classification algorithms J48, Random Tree, Random Forest and AD Tree algorithms are taken and compared with the ETD (Efficient Trim Down) algorithm. The existing decision tree classifications are validated with the 200 datasets also the correctly classified, incorrectly classified, absolute error, mean error and statistical errors are identified and finally compared with the ETD classification algorithm.

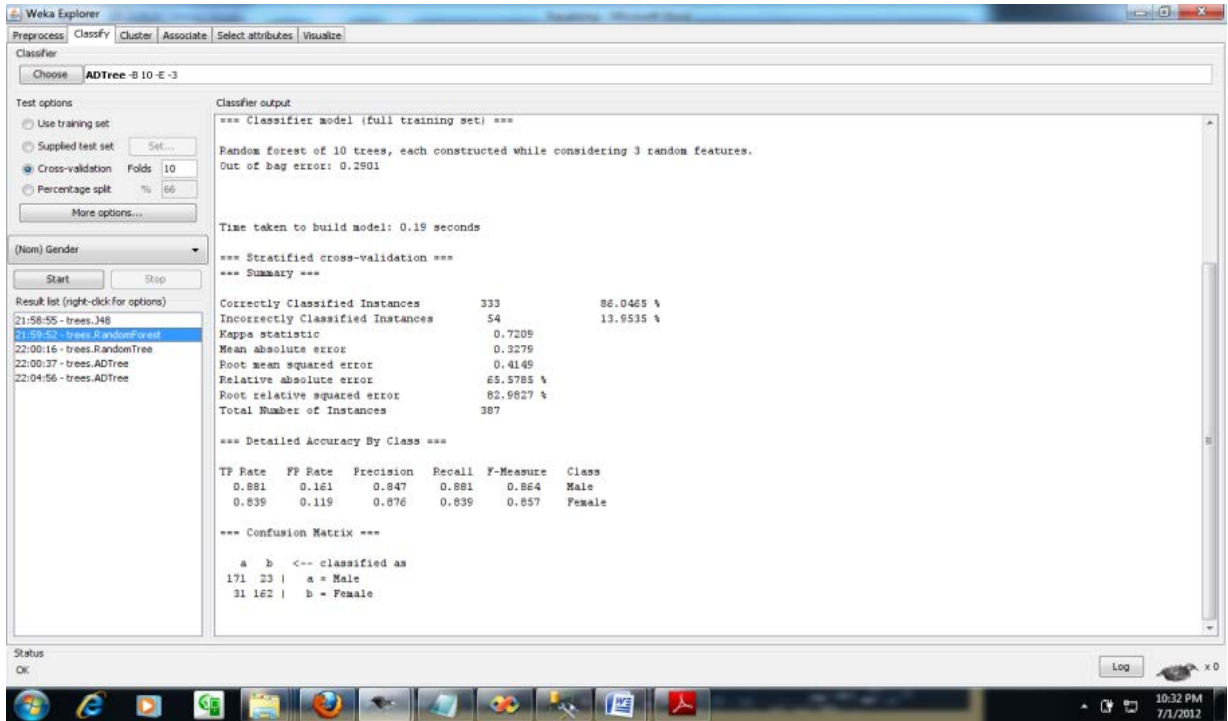
*Different Classifiers (CGP with ETD classifier)*

The below screen shots (4,5,6,7) displays the accuracy and validation of the 200 datasets form the web service transactions for (J48, Random Tree, Random Forest, AD Tree) of the WEKA classifier. The classifier evaluate the data with the standard parameters of Shope code, bill no, card no and amount.

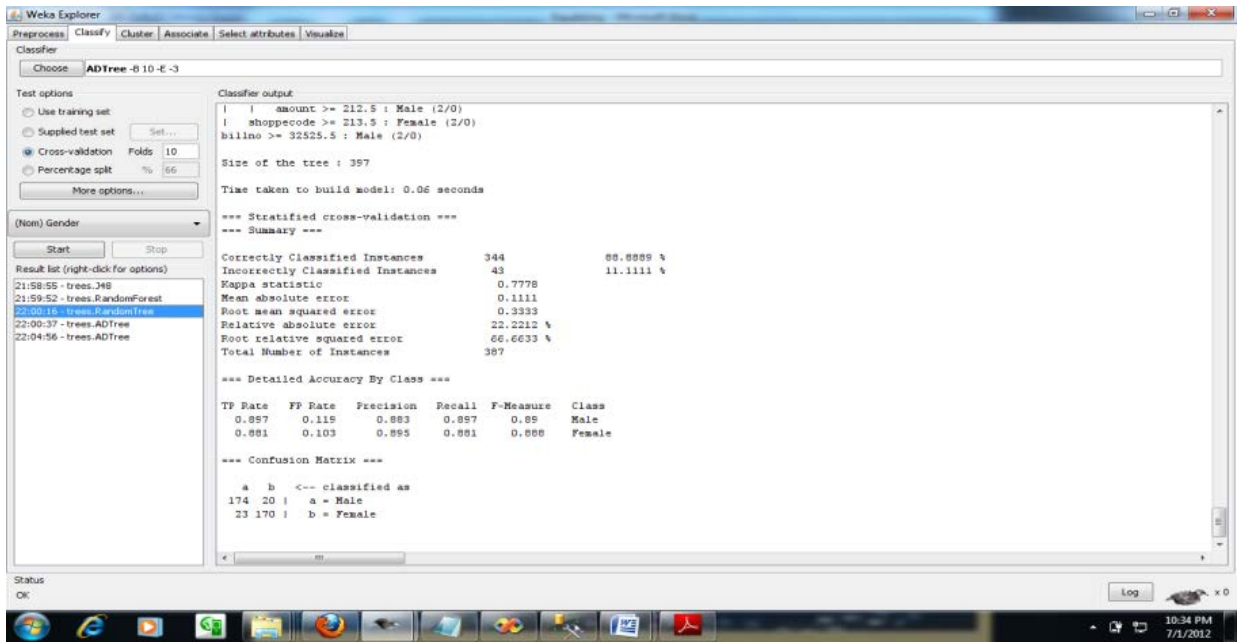
**Screen shot 4: J48**



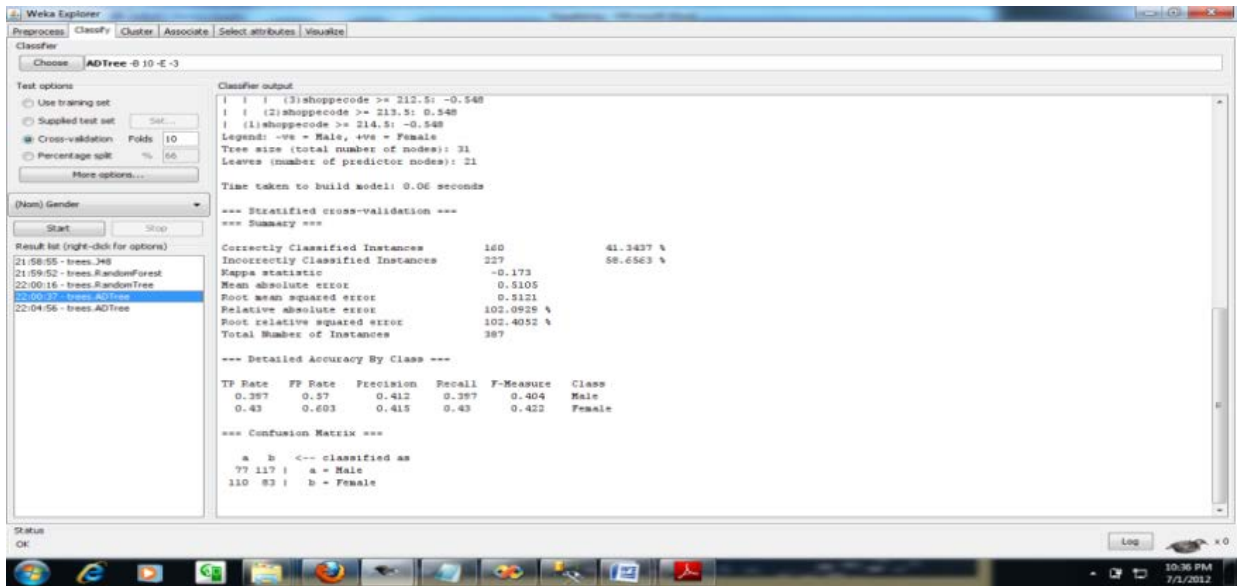
Screen shot 5: Random Forest



### Screen shot 6: Random Tree

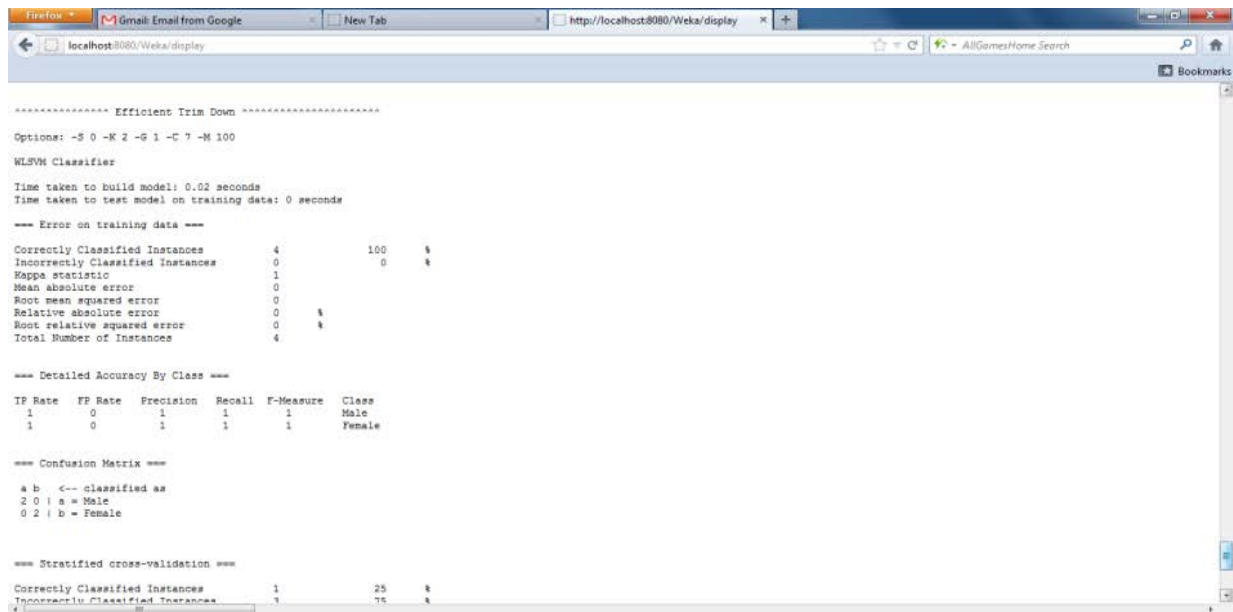


### Screen shot 7: AD Tree



Comparing with the CGP Classifiers the ETD Classifier produces the effective validation and better accuracy. The screen shot.8. displays the resultant output of the ETD classifier.

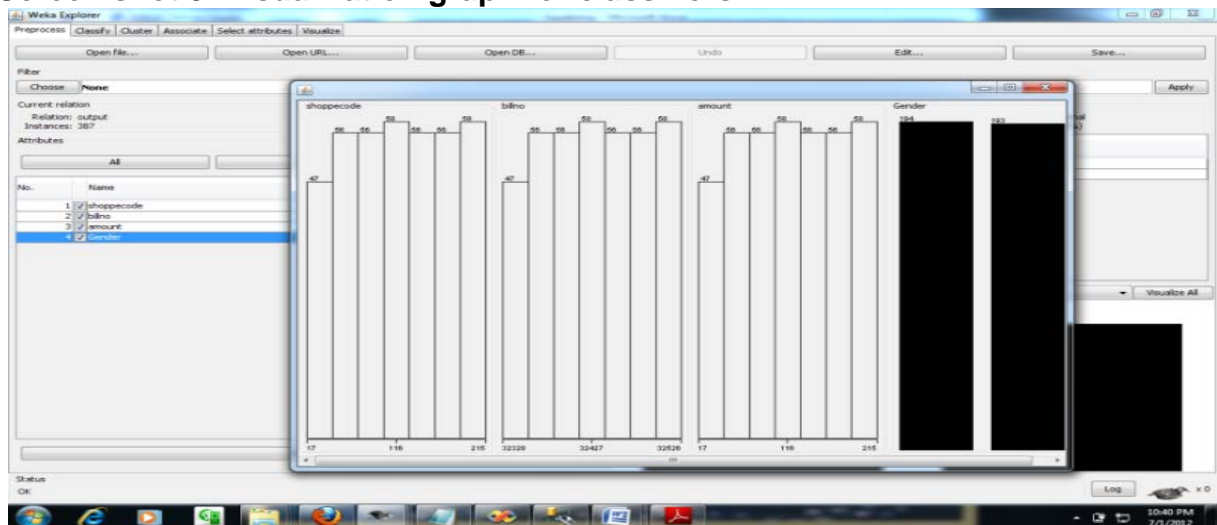
### Screen shot 8: ETD (Efficient Trim Down Classification) Classification



#### Result analysis

The screen shot (9) displays the visualized graph generated with the billno, shope code, amount and Gender. The graph displays the results for J48, Random Tree, Random Forest and AD Tree Classifier. The WEKA Tool visualizes the results for each given classifiers. The black vertical blocks show the male and female category, based on the category the vertical lines displays the values ranges of billno, shope code and amount for the CGP classifiers.

### Screen shot 9. Visualization graph for classifiers





### Stratified Cross Validation and Accuracy measure of (CGP and ETD)

#### Stratified Cross Validation

The Stratified cross validation displays the validation progression of classifiers J48, Random Forest, Random Tree, AD Tree and ETD Classifiers. The correctly classified Instance range and incorrectly classified Instances ranges are stated for the variation of different classifiers with the range of 200 data. To evaluate the validation, the credential attributes assigned are (shope code, bill no, card no and amount). The Correctly classified Instance is 95.5% result (ETD) than compared to the classifiers of J48 (50.1292%), Random Tree (86.0465%), Random Forest (88.8889%) and AD Tree (41.3437%). Also, the validation results shows the set of errors Instances accumulated during the validation.

The given Kappa Statistic is a measure of different categorical items. The group of raters classifies the N items into the entirely different categories. By the Kappa statistic errors for J48 is (0), Random Forest is (0.7209), Random Tree is (0.7778), AD Tree is (0.173) and ETD is (1). Comparing with all these classifiers, the ETD kappa error measure is low with the value of 0.128.

So, the Kappa error measure analysis shows that the categorical items error for ETD is lower than the other classifier analysis. The Root Mean Square error is the measure to quantify, the difference between the estimator values and the true values of the quantity. Comparison of ETD with CGP shows that ETD less error than the CGP classifiers. Similarly, Relative absolute error and root relative square error for ETD are also (0) less when compared with the CGP classifiers. Therefore, the validation defines the tested results of ETD produce a 100 % result compared than the existing classifiers.

**Table.1. displays the validation result of CGP classifiers and ETD classifier**

| Verification                    | CLASSIFIERS |               |             |            |      |
|---------------------------------|-------------|---------------|-------------|------------|------|
|                                 | J48         | Random Forest | Random Tree | AD Tree    | ETD  |
| Correctly classified Instance   | 50.1292 %   | 86.0465 %     | 88.8889 %   | 41.3437 %  | 100% |
| Incorrectly classified Instance | 49.8708 %   | 13.9535 %     | 11.1111 %   | 58.6563 %  | 0    |
| Kappa Statistic                 | 0.128       | 0.7209        | 0.7778      | 0.173      | 0    |
| Mean Absolute Error             | 0.5         | 0.3279        | 0.1111      | 0.5105     | 0    |
| Root mean squared error         | 0.5         | 0.4149        | 0.3333      | 0.5121     | 0    |
| Relative absolute error         | 100 %       | 65.5785 %     | 22.2212 %   | 102.0929 % | 0    |
| Roer relative squared error     | 100 %       | 82.9827 %     | 66.6633 %   | 102.4052 % | 0    |
| Total No of Instance            | 387         | 387           | 387         | 387        | 387  |

*Detailed Accuracy measure of class*

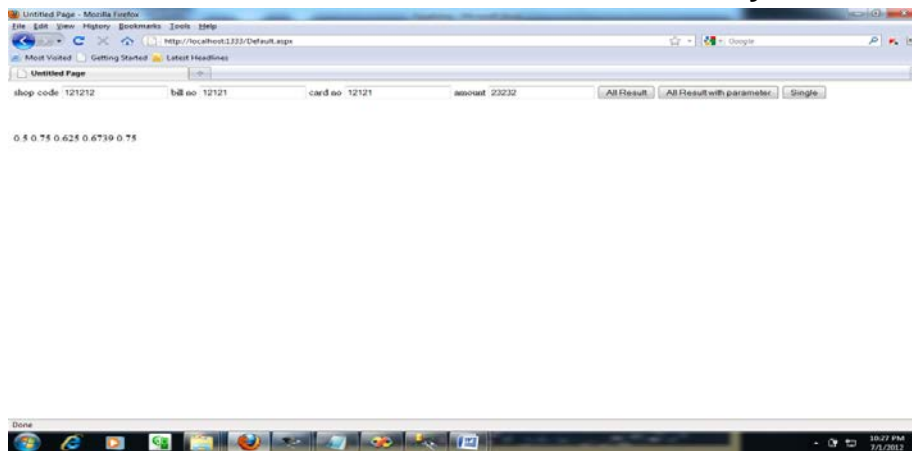
In the detailed accuracy measure of class, the accuracy rate is categorized into Male and Female. Based on the client requests, the accuracy measure is collected a per the gender data set. From Table.2.for bulk data set evolution it can be observed that each classifier produces the less accuracy than the ETD classifier, expect that the FP rate of both male and female category is lower than the other classifiers. So, from the analysis of the data set, ETD produces a better accuracy than the existing CGP classifiers.

**Table.2.displays the accuracy result of CGP classifiers and ETD classifier**

|           | Male   |               |             |         |     | Female |               |             |         |     |
|-----------|--------|---------------|-------------|---------|-----|--------|---------------|-------------|---------|-----|
| Accuracy  | GetJ48 | Random Forest | Random Tree | AD Tree | ETD | GetJ48 | Random Forest | Random Tree | AD Tree | ETD |
| TP Rate   | 0.237  | 0.881         | 0.897       | 0.397   | 1   | 0      | 0.839         | 0.881       | 0.43    | 1   |
| FP Rate   | 0.46   | 0.161         | 0.119       | 0.57    | 0   | 0      | 0.119         | 0.103       | 0.603   | 0   |
| Precision | 0.501  | 0.847         | 0.883       | 0.412   | 1   | 0      | 0.876         | 0.895       | 0.415   | 1   |
| Recall    | 0.432  | 0.881         | 0.897       | 0.397   | 1   | 0      | 0.839         | 0.881       | 0.43    | 1   |
| F-Measure | 0.668  | 0.864         | 0.89        | 0.404   | 1   | 0      | 0.857         | 0.888       | 0.422   | 1   |

Screen shot.9. displays the accurate data generated from WEKA to the web service provider. Here the classified accurate data sent to the Bank processing. If the insufficient accuracy and validate results are not met then the invalid request is again sent to the client.

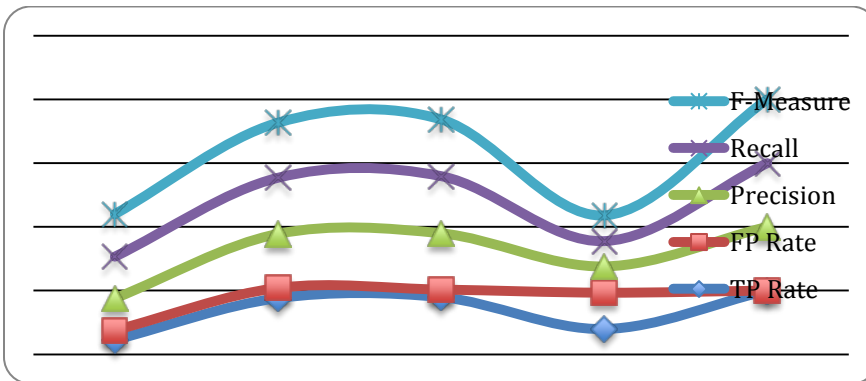
**Screen shot.9.Web service classification accuracy result using EVC**



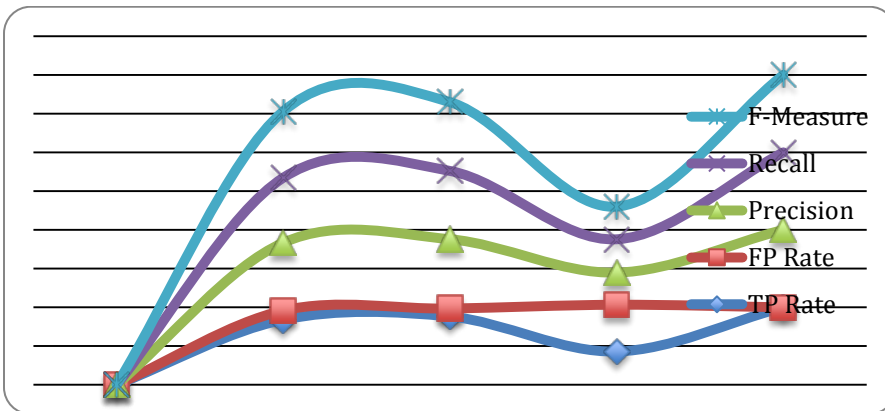
### Performance Analysis of CGP and ETD

The performance analysis of CGP and ETD are shown with the accuracy and validation result of both the CGP and ETD. The dark blue color denotes J48, light green color displays the Random Forest, Violet color displays the Random Tree and the light blue displays the AD Tree. In this comparison, the red colored graph line (ETD) shows more accuracy in terms of FP Rate, Precision, TP Rate and Recall-measure than the CGP classifiers for both Male and Female.

**Accuracy Graph.1. Comparison of CGP and ETD (MALE)**



**Accuracy Graph.2. Comparison of CGP and ETD ( FEMALE)**



### CONCLUSION

In this paper new Classification technique called Equalizing Validation Classification (EVC) method and ETD are proposed for the classification of web service client requests. The accuracy and validation of the vast web service transaction requests are compared with the existing classifiers of (J48, Random Tree, Random Forest and AD Tree) with the ETD (Efficient Trim Down) Classifier. From comparison, it is observed that the proposed ETD (Efficient Trim Down) produces a better accuracy, validation and good response time. In future this work can be extended with the fuzzy technologies to reduce the response time of the requests during the transaction.

## REFERENCES

- Stephen S. Yau, Fellow, IEEE, and Yin Yin (2008), A Privacy Preserving Repository for Data Integration across Data Sharing Services, IEEE Transactions on Services Computing, vol. 1, no. 3, July-September, 130-140.
- Sattar Hashemi, Ying Yang, Zahra Mirzamomen, and Mohammadreza Kangavari. (2009), Adapted One-versus-All Decision Trees for Data Stream Classification, IEEE Transactions on Knowledge And Data Engineering, Vol. 21, No. 5, 624-637.
- Qianxiang Wang, Jin Shao, Fang Deng, Yonggang Liu, Min Li, Jun Han, Hong Mei (2009). An Online Monitoring Approach for Web Service Requirements, IEEE Transactions on Services Computing, Vol. 2, No. 4, 338-351.
- Dimitrios Skoutas, Dimitris Sacharidis, Alkis Simitsis, and Timos Sellis (2010). Ranking and Clustering Web Services Using Multicriteria Dominance Relationships, IEEE Transactions on Services Computing, Vol. 3, No. 3, 163-177.
- Michael von Riegen, Martin Husemann, Stefan Fink, and Norbert Ritter (2010). Rule-Based Coordination of Distributed Web Service Transactions, IEEE Transactions on Services Computing, Vol. 3, No. 1, 60-72.
- Claudio A. Ardagna, Sabrina De Capitani di Vimercati, Stefano Paraboschi, Eros Pedrini, Pierangela Samarati, Mario Verdicchio (2011). Expressive and Deployable Access Control in Open Web Service Applications, IEEE Transactions On Services Computing, Vol. 4, No. 2, 96-109.
- Jung-Yi Jiang, Ren-Jia Liou, and Shie-Jue Lee (2011) A Fuzzy Self-Constructing Feature Clustering Algorithm for Text Classification, IEEE Transactions On Knowledge And Data Engineering, Vol. 23, No. 3, 335-349.
- Nils Gruschka, Meiko Jensen, Luigi Lo Iacono, and Norbert Luttenberger (2011) Server-Side Streaming Processing of WS-Security, IEEE Transactions On Services Computing, Vol. 4, No. 4, 272-282.
- George Zheng and Athman Bouguettaya (2009) Service Mining on the Web, IEEE Transactions On Services Computing, Vol. 2, No. 1, 65-78.
- Fumiko Satoh and Takehiro Tokuda (2011) Security Policy Composition for Composite Web Services, IEEE Transactions On Services Computing, Vol. 4, No. 4, 314-327.
- Salvatore Ruggieri (2002) Efficient C4.5 IEEE Transactions on Knowledge and Data Engineering, Vol. 14, No. 2, 315-344.
- Youakim Badr, Fellow, Frede rique Biennier, Samir Tata, Fellow (2011) The Integration of Corporate Security Strategies in Collaborative Business Processes, IEEE Transactions On Services Computing, Vol. 4, No. 3, 253-254.
- M. Weiss , B. Esfandiari, Y. Luo (2007) Towards a classification of web service feature interactions, ELSEVIER, Computer Networks 51, 359-381.
- Junsup Lee, Sungdeok Cha, Dongkun Leec, Hyungkyu Lee (2009) Classification of web robots: An empirical study based on over one billion requests, ELSEVIER computers & security 28, 795 - 802.
- Grigorios Tsooumakas , Ioannis Vlahavas (2007) An interoperable and scalable Web-based system for classifier sharing and fusion, ELSEVIER Expert Systems with Applications 33, 716-724.
- Nicolás Tourné, Daniela Godoy (2012) Evaluating tag filtering techniques for web resource classification in folksonomies, ELSEVIER Expert Systems with Applications 39, 9723-9729.
- Shuchuan Lo (2008) Web service quality control based on text mining using support vector machine, ELSEVIER Expert Systems with Applications 34, 603-610.

- Eda Baykan, Monika Henzinger, Ludmila Marian, Ingmar Weber (2011)  
A Comprehensive Study of Features and Algorithms for URL-Based Topic  
Classification ACM Transactions on the Web, Vol. 5, No. 3, Article 15.
- Guangyu Zhu, Gilad Mishne (2012) ClickRank: Learning Session-Context Models to  
Enrich Web Search Ranking ACM Transactions on the Web, Vol. 6, No. 1,  
Article 1.
- Radu Jurca, Florent Garcin, Arjun Talwar, Boi Faltings (2010) Reporting Incentives and  
Biases in Online Review Forums, ACM Transactions on the Web, Vol. 4, No. 2,  
Article 5.