

# A Generic Model for Student Data Analytic Web Service (SDAWS)

J. Shana  
Coimbatore Institute of Technology  
Coimbatore, India

T. Venkatachalam  
Coimbatore Institute of Technology  
Coimbatore, India

**Abstract:** Any university management system accumulates a cartload of data and analytics can be applied on it to gather useful information to aid the academic decision making process. This paper is a novel attempt to demonstrate the significance of a data analytic web service in the education domain. This can be integrated with the University Management System or any other application of the university easily. Analytics as a web service offers much benefits over the traditional analysis methods. The web service can be hosted on a web server and accessed over the internet or on to the private cloud of the campus. The data from various courses from different departments can be uploaded and analyzed easily. In this paper we design a web service framework to be used in educational data mining that provide analysis as a service.

**Keywords:** educational data mining, web service, data analytics, association rule mining, classification

## 1. INTRODUCTION

Analytics is a fast emerging area of technology because there is flooding of data due to the advancement in internet. There is huge amount of data everywhere and organizations need to invest into technologies and infrastructure to analyze this data and derive meaningful knowledge out of it for effective decision making process. Education domain also generate lot of academic data that can be mined to gain useful information using which lot of remedial measures can be taken. Of late cloud computing has taken a centre stage for it is able to deliver everything as a service and the organization need no longer buy expensive hardwares or softwares to build the infrastructure needed. And web service is one of the many technologies that is helping cloud computing to achieve its objective. Data analytics involves lot of machine learning and statistical techniques to mine out useful decision making information. Data mining algorithms such as association rule mining, classification clustering and outlier analysis try to bring out hidden meaningful information from huge academic data sets. The web services approach uses standards-based interfaces for connecting data providers with data users. The network strives to go beyond searching and visualizing data to include data processing and analysis services to allow users to create new content. From a technical perspective, the web service technologies consist of a collection of standard protocols that enable the creation, distribution, discovery and integration of software components over the internet. Central to the web service technologies are the concepts of “software as service” and “platform independence” [4].

Analysis software for the education domain differs from mining the financial data or the business data in general. So the traditional commercial analysis softwares cannot satisfy the needs of the institution. This paper suggests a model that implements the analysis techniques as a service based architecture. This would enable the faculty to access the service from a any browser based client application,

irrespective of the language in which the service is built. It can be invoked by any kind of application over the network. This would provide a greater flexibility to integrate even with the legacy client application or the latest kind of mobile application. Also the model can be extended to work on the private cloud of the institution there by leveraging the power of cloud computing.

## 2. THE BACKGROUND

There exists many commercial data analytics service for different domains offering little support for complete data analysis. There is no service for the educational domain to the best of our knowledge. This paper takes the inspiration from such existing attempts. Data analytics as a service is fast catching up after the success of cloud computing. Table 1 gives an overview of companies that have adopted to offer data storage and effective retrieval service commercially. Except ADABA all other products provide only data storage and query support to the clients.

**Table 1: Commercial data analysis softwares**

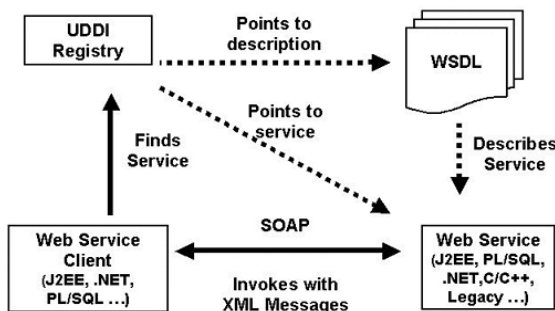
Company	Service	Features
Google	BigQuery	Data analysis for the uploaded data in the cloud.
HP	Vertica Analytic DB	A RDBMS for building data warehouse and database for BI in the cloud
Amazon	SimpleDB	Data storage and retrieval in cloud

Literature survey of the web service analytics gives some interesting information.[6] provides a automatic advice for scientists in their data analysis using data analysis web service.[5] describes the development and application of web service for comparing US and global emission inventories.[8]specifies an on demand BI architecture on cloud for the healthcare sector. [7] offers a wide range of techniques that support the transformation of Data Analysis including web service technology.[9] propose a novel infrastructure that offers complex and optimized query facilities for web services through combined invocation of different web services. Academic analytics is used to derived intelligent information from academic data and used in higher education in decision making. Many notable works have been done in the field of educational data mining. [1] Compares different data mining methods for classifying students based on the Moodle usage data.[3] elaborates on the challenges in educational data mining. [11] uses data mining techniques for identifying the factors that affect the student drop out. In [2] different classification algorithms are implemented to identify the factors that affect the performance of students in a particular course.

### 3. THE PROPOSED SYSTEM

#### 3.1 Web Service Architecture

Fig 1 Web service

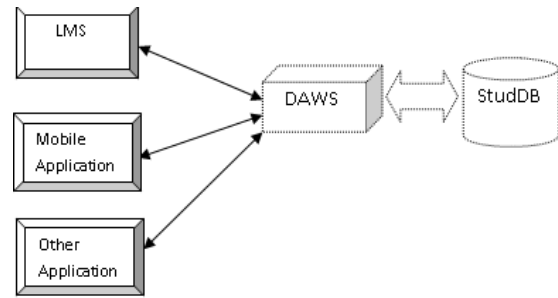


The SOAP-based Web service architecture comprises three entities:

- A service provider, which creates a SOAP based Web service and publishes the service description in the service registry
- A service registry, which enables online service discovery and
- A service requestor, which finds the service by querying the service registry. The requestor then retrieves the service description, uses it to bind to the service implementation, and begins interacting with it

#### 3.2 Architecture of SDAWS

Fig 2: Components of SDAWS



#### Server Side:

**Database:** This holds the entire student data that is represented after efficient preprocessing. The attributes relevant for analysis is selected and stored here.

**Web server:** It here that the data analytical web service resides. This exposes the analysis functionalities to the client application over the network.

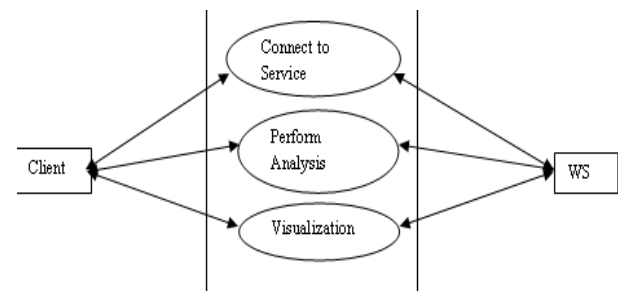
#### Client Side:

This can be any of kind of thin client application that can access the service by connecting to it. The client application can be a part of the Learning Management System (LMS) or any other application running in any of the department of the institution. It can also be a mobile application to be used from anywhere in the campus.

#### 3.3 Web Service Implementation

The web service can be implemented in any language that supports a web service development framework. The client application can be any web based application developed in any language irrespective of the language in which the service is developed. The major functionalities of the data analytic web service are shown as a Use-Case model in Fig 3.

Fig 3: Use case diagram for the proposed system



#### 3.4 Analysis Techniques

The web service can implement any of the data mining functionalities on the student data. The most common mining algorithms that can generate useful information are Association rule mining, Clustering and Prediction techniques as given in [12].

##### Association Rule mining

Association analysis discovers association between the attributes of a relation and specifies it in the form  $X \Rightarrow Y$ .

Such rules are called as association rules. This technique when applied on the student data can reveal useful associations between attributes such as test marks and final result, class attendance and final result and so on. Many algorithms exist for mining association rules and the most noted one is the Apriori algorithm.

#### **Clustering Technique**

Cluster analysis allows objects to be grouped into clusters or groups such that objects within a cluster have high similarity in comparison to one another but very dissimilar to objects in other clusters. This analysis on the student data could help us with information like what kind of students tend to score well in exams and which group of students deviates from the usual. Cluster analysis is helpful for identifying the outliers easily. Algorithms such as K-Means, K-medoids and K nearest neighbor can be used for cluster analysis.

#### **Prediction Methods**

Prediction in data mining involves classification technique to predict the class label of the data objects whose class label is unknown. This method can be applied on the student dataset to predict the performance of the students. [2]

It can also be used to predict the prospective students for a course and to would help in improving the course in general. Decision tree based algorithms such as ID3 or Neural network based algorithms or Bayesian Belief Networks can be used for classification.

#### **4. CONCLUSION**

This paper has given a general model for building a data analytic web service. Analytics as a service provides lot of benefits and can be implemented by any educational institution easily. Once the web service is hosted in the campus server it can be accessed from any browser based application. Academic analysis would help the faculty to determine lot of information about the performance of students in courses and take timely action whenever necessary. This system can be easily incorporated as a cloud based application in the private cloud of the campus also.

#### **5. REFERENCES**

- [1] Romero, C Ventura, S, Espejo, P,G and Hervas, C. 2008. Data Mining Algorithms to Classify Students. *1st International Conference on Educational Data Mining, Proceedings*. Montreal, Canada.
- [2] Shana, J and Venkatachalam, T 2011. Identifying Key Performance Indicators from Student Course Data, *IJCA*.
- [3] Baker, R, J, D and Yacef ,K. 2009 The state of educational data mining in 2009: A review and future visions”, *Journal of Educational Data Mining*.
- [4] Hsing, K , Cheng, Qian and Q, T, Leon ,Z.2006. Web Services and Service-Oriented Application Provisioning: An Analytical Study of Application Service Strategies, *IEEE Transactions on engineering Management*.
- [5] Falke, S, Stella, G and Keating, T. 2007. Web Services for Comparative Data Analysis of Emission Inventories.
- [6] Kollingbaum, M, J, Cai, K and Norman ,T, J 2004. Web Service Support for Scientific Data Analysis.
- [7] Farber, M, Cameron, M , Ellis, C and Sullivan, J.2000. *Massive Data Analytics and the Cloud*, A whitepaper, Booz Allen Hamilton Publications.
- [8] Indrajit, B and Anandhi, R 2011. Healthcare Data Analytics on the cloud, *Online Journal of Health and Allied Sciences*..
- [9] Quzzani, M and Bouguettaya ,A 2004. Efficient Access to Web Service, *IEEE Computer Society*.
- [10] Zheng, G and Bouguettaya, A 2009. Service Mining on the Web, *IEEE Transactions on Services Computing*,
- [11] Ho Yu, C and DiGangi, S. 2010. A Data Mining Approach for Identifying Predictors of Student Retention from Sophomore to Junior Year, *Journal of Data Science*.
- [12] Han, J and Kamber, M 2006. *Data Mining – Concepts and Techniques*, Morgan Kaufmann Publishers.
- [13] <http://www.axis.apache.org/axis2/java/core.htm>