

**School of Information Technology and Engineering (SITE)**

**Master of Computer Application (MCA)**

**Course Project Report**

**Healthcare System Analysis using Cloud Computing**

**Submitted for the Course ITA 6009: Cloud Computing**

**Offered by Dr. R. K. NADESH during WINTER 2018-2019**

***By***

**SABYASACHI CHAKRABORTY** **18MCA0234**

**PRATIK KUMAR DUTTA** **18MCA0203**

**SAHIL GANDHE** **18MCA0207**

**APRIL 2019**



**School of Information Technology & Engineering**

**Master of Computer Application (MCA)**

**ITA 6009: Cloud Computing Winter 2018-2019**

**A Report on the Course Project**

**Healthcare System Analysis using Cloud Computing**

|  |
| --- |
| Team Name : **AZURE MINDS** |
| Team Member(s) with Reg# and Name :  **SABYASACHI CHAKRABORTY; 18MCA0234; 9038786568;** [sabyasachi.chakraborty2018@vitstudent.ac.in](mailto:sabyasachi.chakraborty2018@vitstudent.ac.in)  **PRATIK KUMAR DUTTA; 18MCA0203; 7059652163;** [pratikkumar.dutta2018@vitstudent.ac.in](mailto:pratikkumar.dutta2018@vitstudent.ac.in)  **SAHIL GANDHE; 18MCA0207; 8889391481;**  [sahil.gandhe2018@vitstudent.ac.in](mailto:sahil.gandhe2018@vitstudent.ac.in) |
| Project Title : Healthcare System Analysis using Cloud Computing |
| 1. **Introduction**    1. Background (System Study Details in brief)   *Distributed cloud computing contains mobile and cooperative applications and administrations. Expanded capacity with high computerization, adaptability and less expense are a couple of positive circumstances distributed cloud computing offers. Utilizing distributed cloud computing can develop social insurance administrations transference for patients. This paper describes the classifications and administration models of distributed cloud computing system and specially its innovative knowledge, Data mining classification techniques for huge data analysis in medicinal services. Issues identified with cloud security, boundaries to cloud systems and consistency with acts of accuracy in medicinal services were additionally exhibited.*  ***Keywords- Health care System Analysis learning, Cloud Storage platform, Communication and Networking, Usage of Data Mining classifications.***   * 1. Problem Statement   An authorized user can get the detailed percentage of diseases from dataset around the country for health care analysis of diseases stored in cloud platform. It can accessible from anywhere. User will be able to get faster results any time from anywhere as it is implemented in cloud platform.  Admin should gain the access to log modify, add and manipulate data along with providing a general accuracy rate by using Data Mining algorithms and perform the process in cloud based platform.   * 1. Importance   Healthcare System Analysis and storing in cloud platform is the task of identifying and categorizing diseases expressed in a statement, especially in order to determine whether the result of a patient is curable or non-curable, and storing it through a network in virtualized storage platform which can be accessible from anywhere, anytime. |
| 1. **Overview and Planning**    1. Proposed System Overview  * OS Version of Windows 10 * i5 processor * 4GB RAM * Operating system 32x or 64x bit * Cloud platform (Amazon Web Services) * Python terminal (Cloud9) * 1 TB Hard drive.   1. Challenges * Cloud storage is vast enough to store large dataset of patients, staff and doctors. * Backup dataset should always be prepared for the times when dataset crash happens. * Prediction of diseases whether it’s curable or not for society care. * Cloud platform security. * Convenient and easy accessibility.   1. Assumptions   In this project we’re studying about Naive Bayesian Classification, Random Forest, K-Means Clustering, SVM and other algorithms and researched of that paper. We have selected one primary dataset to use in Algorithms and code to see the efficiency and accuracy in Amazon web service platform. After that team has to implement algorithms in programming language Python, import and read in AWS for storing in virtualized OS. Further steps are, executing accurately and getting an approximate result we are going to compare algorithms to check the most efficient option and provide better accuracy of health care analysis for diseases.  2.4 Architecture Specifications     * Admin can access AWS and add modify dataset. * Architecture is fully based in Cloud forum and accessed from anywhere * User from one another physical layer access virtual storage web application designed by one admin and access the records and other concerns. * Cloud architecture can read dataset and implement algorithms for results accuracy. * Server app architecture is interactive for user so that they can access their own queries from one single platform.   The architectural diagram as follows:  (Ellipse denoted as users, Rectangles are processes)     * Minimum of 20GB Hard Drive space, * Min of 2 GB RAM, * Resolution 1024 X 768 or higher   1. **Software Requirements:** * Windows 8 or 10, * i3 or i5 processor 2 GHz or more, * 4GB RAM, * Operating system 32x or 64x bit |
| 1. **System Design**   3.1 High-Level Design   * Admin can access AWS and add modify dataset. * User from one another physical layer access virtual storage web application designed by one admin and access the records and other concerns. * Cloud platform can read dataset and implement algorithms for results accuracy. * Web application interface for user so that they can access their own queries from one single platform.         3.2 Low-Level Design  Use case diagram showing levels of accessible components by user and admin |
| 1. **System Implementation**   4.1 Module Development –Code  Dataset:    Code:  import pandas as pd  import numpy as np  from sklearn.naive\_bayes import BernoulliNB  #import seaborn as sns  #import matplotlib.pyplot as plt  #% matplotlib inline  data=pd.read\_csv("DiabetesData.csv")  print(data.head(10))  print("# no of passenger in the data set:", +(len(data)))  ##Analysis data  #sns.countplot(x="diastolic bp",data=data)  #sns.countplot(x="diastolic bp",hue="age", data=data)  #plt.show()  # sns.countplot(x="loans",hue="homeowner",data=bank\_data)  # plt.show()  #data["age"].plot.hist()  #plt.show()  #  ## TRAIN MY DATASET  x=data.drop("class",axis=1)  y=data["class"]  from sklearn.model\_selection import train\_test\_split  x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.2,random\_state=1)  from sklearn.linear\_model import LogisticRegression  logmodel=LogisticRegression()  print(logmodel.fit(x\_train,y\_train))  predictions=logmodel.predict(x\_test)  from sklearn.metrics import classification\_report  print(classification\_report(y\_test,predictions)) #generate classification report  ##generate accrucy  from sklearn.metrics import confusion\_matrix  print(confusion\_matrix(y\_test,predictions))  from sklearn.metrics import accuracy\_score  print( "The Accuracy of the Logistic Regression prediction is: ",accuracy\_score(y\_test,predictions))  x\_train, x\_test, y\_train, y\_test = train\_test\_split(x, y, test\_size=0.2, random\_state=1)  from sklearn.tree import DecisionTreeClassifier  classifier = DecisionTreeClassifier()  classifier.fit(x\_train, y\_train)  y\_pred = classifier.predict(x\_test)  from sklearn.metrics import classification\_report, confusion\_matrix  print(confusion\_matrix(y\_test, y\_pred))  print(classification\_report(y\_test, y\_pred))  from sklearn.metrics import accuracy\_score  print("\n======================================================================================")  print( "The Accuracy of the prediction using Decision Tree is: ", accuracy\_score(y\_test, y\_pred))  x\_train, x\_test, y\_train, y\_test = train\_test\_split(x, y, test\_size=0.35,random\_state=1)  BernNB = BernoulliNB(binarize= 0.1)  BernNB.fit(x\_train,y\_train)  print(BernNB)  y\_expt = y\_test  y\_pred = BernNB.predict(x\_test)  print("\n======================================================================================")  print("The Accuracy using Naive Bayes is: ",accuracy\_score(y\_expt, y\_pred))  4.2 Output/Results  The followings are the output of implementation in AWS cloud9            4.3 Discussion   * Reading datasets are in .csv extension and creating storage for that in S3 of Amazon Web Services with azmapplication named environment. * Creating instances and virtualized OS in EC2 in AWS platform. * Creating Web app and deploying self modified HTML code in Elastic beanstalk of AWS. * Implementation of code in Cloud9 platform by Python 3 or 2 or uploading using external terminal like py charm in AWS. * Reading code from physical desktop to virtual machine and reading to cloud9 compiler in AWS.  1. Conclusion and Future Developments   In this paper, Health care analysis is performed by collecting the dataset and individual pacifiers such as three algorithms mentioned in the project (Logistic regression, Decision tree, and Naïve Bayesian classifier).  The individual classifier accuracy of 77% is obtain for one classifiers which is Logistic Regression and for decision tree it is 68% and for Bayesian Classifier accuracy is less such as 61% however in the future, an ensemble of three algorithms is deployed for better accuracy.   1. References   *[1] Zhang, Yin, et al. "Health-CPS: Healthcare cyber-physical system assisted by cloud and big data." IEEE Systems Journal 11.1 (2017): 88-95.*  *[2] Zhang, Y., Qiu, M., Tsai, C. W., Hassan, M. M., & Alamri, A. (2017). Health-CPS: Healthcare cyber-physical system assisted by cloud and big data. IEEE Systems Journal, 11(1), 88-95.*  *[3] Zhang, Y., Qiu, M., Tsai, C.W., Hassan, M.M. and Alamri, A., 2017. Health-CPS: Healthcare cyber-physical system assisted by cloud and big data. IEEE Systems Journal, 11(1), pp.88-95.*  *[4] Hassanalieragh, Moeen, et al. "Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: Opportunities and challenges." 2015 IEEE International Conference on Services Computing. IEEE, 2015.*  *[5] Lo’ai, A. Tawalbeh, et al. "Mobile cloud computing model and big data analysis for healthcare applications." IEEE Access 4 (2016): 6171-6180.* |
|  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | | **School of Information Technology & Engineering**  **Master of Computer Application (MCA)**  **ITA 6009 : Cloud Computing Winter 2018-2019**  **Course Project- Implementation Review(Final )**  **Evaluation Sheet** | | | |
| **Title: Healthcare System Analysis using Cloud Computing** | | |  | | | |
| **Team Name:** AZURE MINDS | | |  | | | |
| **Project Team** | | | | | | |
| **S.No** | **Register Number** | **Student Name** | | | **Signature** | **Guided By** |
| **1.** | **18MCA0234** | **SABYASACHI CHAKRABORTY** | | |  | **Dr. R.K.Nadesh** |
| **2.** | **18MCA0203** | **PRATIK KUMAR DUTTA** | | |  |
| **3.** | **18MCA0207** | **SAHIL GANDHE** | | |  |
| **Team Member(s) Contribution and Performance Assessment** | | | | | | |
| **Components** | | | | **Student 1** | **Student 2** | **Student 3** |
| **Analysis Component (15)** | | | |  |  |  |
| **Deploying in Cloud(15)** | | | |  |  |  |
| **Completion of project(10)** | | | |  |  |  |
| **Upload and formatting of the report(10)** | | | |  |  |  |
|  | | | |  |  |  |
| **Student Feedback(Student Experience in this Course Project)** | | | | | **Evaluator Comments** | |
| **Name & Signature of the Evaluator(s)** | | | | | | |