ISSN: 2320-2882

IJCRT.ORG



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A CLINICAL DECISION SUPPORT FRAMEWORK FOR HETEROGENEOUS DATA SOURCES

Dr. S. Vijayaragavan, Professor^{1,} Harikrishna kk, Student². Department of computer Science and Engineering, Muthayammal engineering college,Tamilnadu,India

ABSTRACT

To keep pace with the updates in obliging informatics, thriving recuperating data is being amassed unendingly. Regardless, inferable from the not too appalling gathering of its classes and sources, therapeutic data has ended up being particularly jumbled in various specialist's work environments that it now needs Clinical Decision Support (CDS) system for its affiliation. To sensibly utilize the party flourishing data, we propose a CDS structure that can interweave mixed thriving data from different sources, for instance, take a gander at office test works out as planned, critical information of patients, and achievement records into a joined depiction of features everything considered. Using the electronic prospering healing data so made, multiname delineation was used to endorse a layout of afflictions and thusly help experts in diagnosing or treating their patients' therapeutic issues more capably. Once the ace sees the contamination of a patient, the running with organize is to consider the conceivable complexities of that perplexity, which can impel more infections.

Keywords : Clinical Decision Support, analysing and file.

OBJECTIVE:

In this paper we are analyzing clinical data by using hadoop tool along with some hadoop ecosystems like hdfs, mapreduce, sqoop, hive and pig. By using these tools, we can process no limitation of data, no data lost problem, we can get high throughput, maintenance cost also very less and it is an open-source software, it is compatible on all the platforms since it is Java based. In clinical data details, this is based on whenever patient enters their id in hospital database within no time complex their data will present.

INTRODUCTION

It is broadly acknowledged that wellbeing data devices and machine learning methods can be misused effectively to help specialists in diagnosing and treating their patients all the more productively. Utilizing their experience and information, the doctors order patients and analyze their infections, however in doing as such, it is plausible that they submit a few errors, especially when they need sufficient experience or when their staff of judgment is poor. In such circumstances, Clinical Decision Support (CDS) frameworks, including frameworks that give customized restorative estimation, finding, treatment and important learning, would be useful to the doctors by method for giving them particular learning, patients' data and keen applications, which can enhance the proficiency of their basic leadership forms. Cds frameworks center around separating attributes of patients, in light of which they order patients and give relating clinical recommendations to the doctors. Patients' restorative data is restored from their own medicinal information.

EXISTING SYSTEM:

Existing concept deals with providing backend by using mysql which contains lot of drawbacks i.e.,

www.ijcrt.org

data limitation is that processing time is high when the data is huge and once data is lost, we cannot recover so thus we proposing concept by using Hadoop framework.

DRAWBACKS IN EXISTING SYSTEM:

- \checkmark We can process limitation of data.
- ✓ We get results with take more time and miniatous cost is very high.

PROPOSED SYSTEM

In this paper we are analyzing clinical data by using hadoop framework along with some hadoop ecosystems like hdfs, mapreduce, sqoop, hive and pig. By using these tools we can process no limitation of data, no data lost problem, we can get high throughput, maintence cost also very less and it is a opensouresoftware, it is compatible on all the platforms since it is Java based

METHODOLOGIES

Following modules involves

MODULES

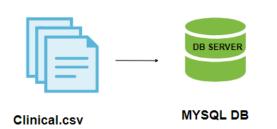
- Preprocessing clinical database
- Storage
- Analyse query
- Analysis latin script (pig)
- Processing (mapreduce)

MODULE DIAGRAMS DESCRIPTION & MODULE

PREPROCESSING CLINICAL DATABASE:

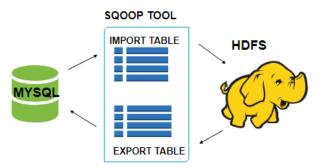
In this module, analyzing the data with different kinds of fields in Microsoft Excel then it converted into comma delimited format which is said to be csv(comma separator value) file and moved to mysql backup through Database.

PREPROCESSING:

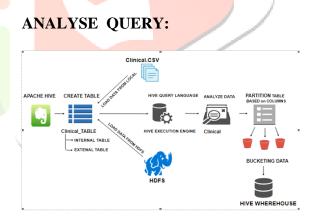


Here by Getting historical data we have to convert those historical batch processing data from (.xlsc) format to (.csv) format and by taking backup of all those data in MYSQL Database to avoid loss of data.

STORAGE:



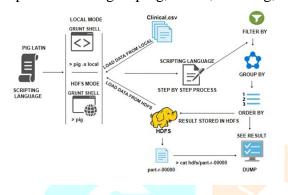
In this module we are getting all those backup data which we have stored in MYSQL and importing all those data by use of sqoop commands to HDFS(Hadoop Distributed File System).now all the data are stored in HDFS were it is ready to get processed by use of hive.



In this module we are getting all those data from HDFS to HIVE by use of sqoop import command .were hive is ready to analyze .here in HIVE we can process only structured data to analyze. by extracting only the meaningful data and neglecting unclenched data we can analyze the data in more effective manner by use of hive.

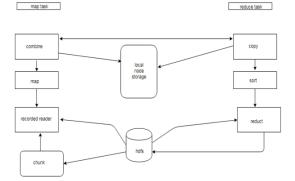
ANALYSIS LATIN SCRIPT (PIG):

To analyze Clincal using Pig, programmers need to write scripts using Pig Latin language and execute them in interactive mode using the Grunt shell. All these scripts are internally converted to Map and Reduce tasks.After invoking the Grunt shell, you can run your Pig scripts in the STORE. shell.Except LOAD and while performing all other operations, Pig Latin statements take a relation as input and produce another relation as output. As soon as you enter a Load statement in the Grunt shell, its semantic checking will be carried out. To see the contents of the schema, you need to use the Dump operator. Only after performing the dump operation, the MapReduce job for loading the data into the file system will be carried out. Pig provides many built-in operators to support data operations like grouping, filters, ordering, etc.



PROCESSING (MAPREDUCE):

Map Reduce is a framework using which we can write applications to process huge amounts of Clinical, in parallel, on large clusters of commodity hardware in a reliable manner. Map Reduce is a processing technique and a program model for distributed computing based on java. The Map Reduce algorithm contains two important tasks, namely Map and Reduce. Map Reduce program executes in three stages, namely map stage, shuffle stage, and reduce stage. The map or mapper's job is to process the input data. Generally the input data is in the form of file or directory and is stored in the Hadoop file system (HDFS). The input file is passed to the mapper function line by line. The mapper processes the data and creates several small chunks of data. This stage is the combination of the Shuffle stage and the Reduce stage. The Reducer's job is to process the data that comes from the mapper. After processing, it produces a new set of output, which will be stored in the HDFS.



SYSTEM TECHNIQUES:

MapReduce is a processing technique and a program model for distributed computing based on java. The MapReduce algorithm contains two important tasks, namely Map and Reduce. Map takes a set of data and converts it into another set of data, where individual elements are broken down into tuples (key/value pairs). Secondly, reduce task, which takes the output from a map as an input and combines those data tuples into a smaller set of tuples. As the sequence of the name MapReduce implies, the reduce task is always performed after the map job.

The major advantage of MapReduce is that it is easy to scale data processing over multiple computing nodes. Under the MapReduce model, the data processing primitives are called mappers and reducers. Decomposing a data processing application into mappers and reducers is sometimes nontrivial. But, once we write an application in the MapReduce form, scaling the application to run over hundreds, thousands, or even tens of thousands of machines in a cluster is merely a configuration change. This simple scalability is what has attracted many programmers to use the MapReduce model.

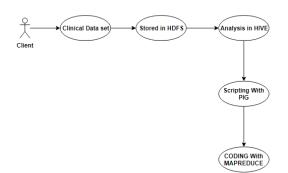
THE ALGORITHM

- Generally MapReduce paradigm is based on sending the computer to where the data resides!
- MapReduce program executes in three stages, namely map stage, shuffle stage, and reduce stage.

Map stage: The map or mapper's job is to process the input data. Generally the input data is in the form of file or directory and is stored in the Hadoop file system (HDFS). The input file is passed to the mapper function line by line. The mapper processes the data and creates several small chunks of data.

Reduce stage : This stage is the combination of the **Shuffle** stage and the **Reduce** stage. The Reducer's job is to process the data that comes from the mapper. After processing, it produces a new set of output, which will be stored in the HDFS.

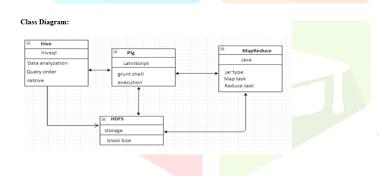
USE CASE DIAGRAM:



EXPLANATION:

A use case diagram is a graphic depiction of the interactions among the elements of a system. A use case is a methodology used in system analysis to identify, clarify, and organize system requirements. In our use case diagram we analyze of clinical data, which will be easy to predicate. Let us see the process of use case diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive,pig and mapreduce.

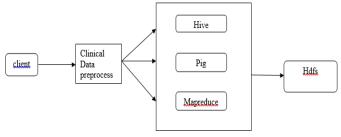
CLASS DIAGRAM:



EXPLANATION:

illustration A class diagram is an of the relationships and source code dependencies among classes in the Unified Modeling Language (UML). In our class diagram the user login to his/her own account. In our class diagram the client can be analysis the clinical, which will be easy to predicate. Let us see the process of use class diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive,pig and mapreduce.

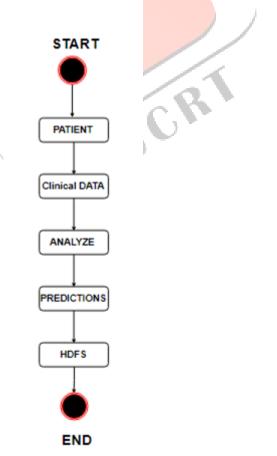
OBJECT DIAGRAM ANALYSIS STORE



EXPLANATION:

An object diagram in the Unified Modeling Language (UML) is a diagram that shows a complete or partial view of the structure of a modeled system at a specific time. If the user is new register the account. In object diagram the client can be analysis the clinical, which will be easy to predicate. Let us see the process of use object diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive,pig and mapreduce.

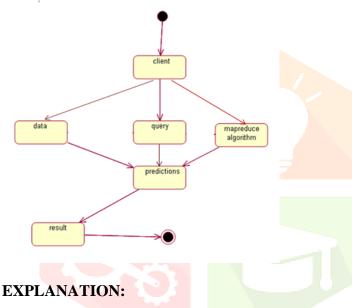
STATE DIAGRAM:



EXPLANATION:

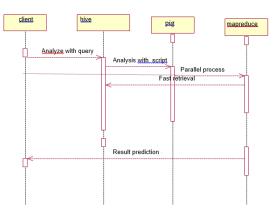
A state diagram is a type of diagram used in computer science and related fields to describe the behavior of systems. State diagrams require that the system described is composed of a finite number of states; sometimes, this is indeed the case, while at other times this is a reasonable abstraction. In state diagram the client can be analysis the clinical, which will be easy to predicate. Let us see the process of use state diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive,pig and mapreduce and call detail record analysis can be delivered into client.

ACTIVITY DIAGRAM:



Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In Unified Modeling Language, activity the diagrams are intended model to both computational and organizational processes (i.e. workflows). We can analysis the clinical, which will be easy to predicate. Let us see the process of use activity diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive, pig and mapreduce and cdranalysis can be delivered into client.

SEQUENCE DIAGRAM:



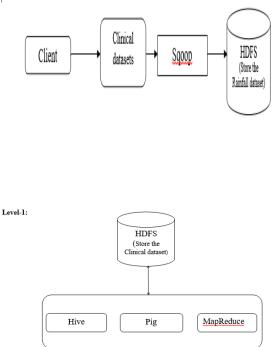
EXPLANATION:

A Sequence diagram is an interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. A sequence diagram shows object interactions arranged in time sequence. In sequence diagram the client can be analysis theclinical data, which will be easy to predicate. Let us see the process of use sequence diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive,pig and mapreduce and power plant analysis can be delivered into client.

D<mark>ATA FLOW DI</mark>AGRAM:

LEVEL-0:

Level-0:

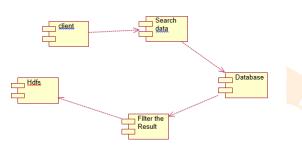


www.ijcrt.org

EXPLANATION:

A Data Flow Diagram (DFD) is a graphical representation of the "flow" of data through an information system, modeling its process aspects. A DFD is often used as a preliminary step to create an overview of the system, which can later be elaborated. In data flow diagram the client can be analysis theclinical data, which will be easy to predicate. Let us see the process of use data flow diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive,pig and mapreduce and Power plant analysis can be delivered into client.

COMPONENT DIAGRAM



EXPLANATION:

In the Unified Modeling Language, a component diagram depicts how components are wired together to form larger components and or software systems. They are used to illustrate the structure of arbitrarily complex systems. In component diagram the client can be analysis theclinical data, which will be easy to predicate. Let us see the process of use component diagram in which whatever data analyzation done to be stores in hdfs (hadoop distribution file system) by different type of styles like hive.pig and mapreduce.

GENERAL

These are the requirements for doing the project. Without using these tools and software's we can't do the project. So we have two requirements to do the project. They are

- 1. Hardware Requirements.
- 2. Software Requirements.

1. HARDWARE REQUIREMENTS

MapReduce The har may serve as the basis for a contract for the implementation of the system and should therefore be a complete and consistent specification of the whole system. They are used by software engineers as the starting point for the system design. It shows what the system does and not how it should be implemented.

PROCESSOR : PENTIUM IV 2.6 Hz.Intel Core 2 Duo.

RAM : 4GB DD RAM MONITOR : 15" COLOR HARD DISK : 40 GB

2 .SOFTWARE REQUIREMENTS

The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is a set of what the system should do rather than how it should do it. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating cost, planning team activities. performing tasks and tracking the teams and tracking the team's progress throughout the development activity.

	Framework	: hadoop
-	Database	: my sql 5.5
	Lanuage	: pig, hive, corejava
_	Data access to	ool: sqoop
	Operating sys	tem: cent os
	Ide	: eclipse

IMPLEMENTATION

GENERAL

In this we implement the coding part using eclipse. Below are the coding's that are used to generate the domain module for electronic text books. Here the proposed techniques are used in the coding part to generate the e-books.

EXISTING APPLICATION (MYSQL):

Dataset Creation

	Cel Tame			· · · · · · ·	2-	the lines			ALL MARKED THE TAXABLE PARTY.	har da
5	Case -			1	1.0	410 LP44			12 P P C C D ger	25 23
	Instan \$11.		der.	A	1. 2.	and dates in 1 + 1	6.4.15			Getty And &
								11		FRAT SINCT
									5m 56 6	
	1 - 5	80.00								
-			-				1.14	-		
100	A COMAND OF		ñ 4	ACCOUNT AND	a in	a man in	and and	10.0	Presi ARDAT UNION LA TRI	HEIGHTS MALMAN
	the service state have been		inite .	DO AMONDO	100	10.00			US DEDG MUNICIPAL ACTIVITY OF A DESCRIPTION	INCOME MATCHED
	tor alteracions. Cre	104	Mer .	full tax, within		MAXING .	10.10		THE DEDUCT MEANING HIS REPORT OF THE ADDRESS OF THE	BM SCOTT HEATSARD
	NULAS-MODEL DA	100	144	Max appendix a	46.00	101044-2184	10.31		NO DECKY DETAILS? IS ANY ANY INSTITUTE A.D.	NW YOOCK REATONNESS
			det -	NV and	100	ATTACK		- e	KE DED THE TANK A BAR	NW TOOLS MATCHING
	ON A JUNKERSKIED MOR		int -	ALC: NOTION TOOL		CARDIN	10.00		TO DEDITITE AND ANALY INCOME.	and tubout meatonems
	DOL INCIDENT CONTINUE OF		011	218 404,44046	1.8	3.64	19.10	18	12 DEDG NEWINI ALERO 1071	SIN' DOOCH NEATHARDS
	DR KIMP-RUSS ENDPLOY	36A 1	100	THE ALMANDER	190	LESYNCLIFE.			10 0400 14/4/0/ KP4//CORDINATE	THE DOCK HEATSAND
			Set .	201404-001	6.4	16004	化用		NE DEDISTRIATION ARE ARE	WW TOOCT, BEATSALIEN
	DUE APOLICIDADESHIDE CH		016	156 KOHH6	2.8	PAUMA.	16 16	- 62	NE CROWN NOTATION ANALON TO FOLLOWING THE	NW TOOCT, REATOALTIN
	SIN WELLSHOWALT DR		CYE	105 40487	15.80	1800	11 11		12 DEDITFERENT MITLEUM MITLEUM	WHY TOOKET, THEATOMATINE
	15 WOLDHOPFALDHOM		100	NO ADD MILLE	41.00	MOR-CELLINDER	25 18		DE DEDIS NETWORK KARANAKE AND TAKEN AND TAKE AST	30V TOOOT T TRATSANDA
	E1493-24064-58504		19.95	105.407-66	41.0	HEARING .	15 10		V5 24045 NEW WAREN SCHEDWITCH MONOCHOD	304 YOGGY HEATSAULIN
			2016	20.8148.44	37	HACALTRACK-R	10 20		NUMBER OF THE PRESENCE OF THE	IEN/100023/THEATUNESIS
			4572	170 108.47181	0.0	HING DRUCK!	10.16	11	to peop televiter account w	IEM TOOCT, TRATITIANS
			100	21014024048	417	ACUERMENCE/PC	17.10		NE DESIS NEARINI NUCEDUVIDE	WY NOCH REALIZABLE
			04	152 100594-0-4816	25.4	AUTOMOREC	. 10, 10,		# 04043 MERARDI M/081198	NW TOOCK MATCHARM
	REAL PROPERTY.		116	URL KOPPORT	7.8	AUEL/HHORAD	10.10	- 81	ID ORDIT MEARINE N. DD JALANTOO ANA	NH COLT MARIABLE
	BIT BOT KOPOL MEAT OR		P75	210 RECEPTION	N.M.	AC/ERIODUCAI	. 10 . 10		NO DEDG NEARINA MEAN CHILDRING COM	WW TOOL'S MEANINGS
			101	549 HOOPField	42.8	AURMEDEDIN	11.14		NU DRIVISTNERARIST SHEARING PROTOKOPE	XN/10001174841344108
			404	10 HORIN	14.94	ASSAULTEN ASSAULTEN	* *		IN DROVE NEARING LAUGH CA.	INV TO OCT. NEATOAANNA
	BOI INVESTIGATION OF THE		101	THE HOLEDING	28	AND ACTES ARE A	0.0		INF DEDISTREAMENT CARGACINE IN LIKE	UNTERSTEINER
								. 5.	NO DEDISTREAMENT CR-15UM	TEM TOXX T. TRAFLACTER
	ELIONICIONA ON		1111	201342414	318	S-CELIFIC OFFICE SAMPTS OFFICE	2.5	- 5	SA DEDIS NEARSY DEBIN SORDA	XM TOOLY REATINGS XM TOOLY RESTORED
				TO MULTINE	- 58	CHOICE HENCET	2.2		THE DECK THE MERINE DISCHARTED IN THE PROPERTY OF THE PROPERTY	HIM YOOLY TRATINGS
	ET LUXE REALTING		<u>.</u>	172 Mar 480	110	OPENERALIZEDE	- 5 - 5	- 2-	THE DECK THE ARRAY DOLLATING, WE THEN CHED.	WW TOOCT TRATINATIN
	EX DADUMENTOPIA OF			TT NUMBER	100	CHORE MILLIONE	2.4	- 2-	VEDEXE NOVEMBER DOLLETING, NET YOU CONST.	INTERNATION CONTRACTOR
	62 Departus, Infrancia Col		100	116 Inc. Alligne	- 22	OPTIMUMENTS.	4.2	- 2-	the Defined free Address of the Definition of the Control of the C	WHY TO COLD THE ATTINETY
	ALL DRITIAL DRIVE DR		and the	CO MUNTY	- 22	YOLGIALINEY	- 2-2	- 2	a DEDUTIONARY COMMITTY	NW TOXXX REPORTS
			22	102 100,0710	- 22	HART CELLUDICEN	2.2	- 2-	WE DROY THATTAC DUSTING INCOMPANY AND COUNTRY LITTINGST	HINT TOXY T TRATIANTIN
	SIT OF THE PERIOD		1018	THE MULTINE	- 22	ADVICTOR OF LCC		- 2	TO DESCRIPTION OF THE REAL PROPERTY OF THE PRO	IN TOXAL MATURA IN
	EN COMPACING AND AND		100	ALC: NO. ADVING	- 22	HOLDING THE HOLDING	- 2- 2	- 21	In priori belanno burrow	EM TODOXY REATONNESS
	EU ORION-ORIA YOM		201	ALC NULLEY AND	6.10	NUMBER OF COLUMN AND A	-2-2	- 2-	NO DEDUCTION AND DATION OF THE MAKE HE	WAY NOOCH TRATINGTON
			WT.	110.100-0.01		ATTRACTORY AND	18-2	- 0-	NO DEDIS NOVATIN' SCREWIGS	UNITION TOWNTOWN
			100	The ball then		LARGE COMMANDER	2.2	- 21	TO OFFICIATION THE AREA THE AREA	MM TONY & MANTANAMA
	en Demokaliseros		Meri	site techner	14.0	LINESCELATING TO		1.22	NE DECKI NE SAMON NE, DALETTING	Star Doorn Searchamp
64	Menth / Mustal / Brand		-		-		100		1	
										0.2
-			-			-				State of the Owner
				C 0 1						· · · · · · · · · · · · · · · · · · ·

Restore Backup File into Mysql Database

_____ [training@localhost ~]\$ mysql -u training -p movielens</home/training/mobility

|---->Database Name Enter password:training



How To Open Mysql Through Terminal

[training@localhost~]\$ ~]\$mysql -u training -p

[training@localhost

Enter password:

Welcome to the MySQL monitor. Commands end with ; or \g . Your MySQL connection id is 7 Server version: 5.0.77 Source distribution

Type 'help;' or '\h' for help.Type '\c' to clear the buffer.

mysql	>
-------	---

mysql> show databases;
++
Database
++
information_schema
movielens
training
++
3 rows in set (0.03 sec)

mysql> use movielens;

Onuberchaining, VMU,2 - MAuser Workstation		0		090
the Balk View VM Tales Help				
Codes, James, M., 27 X Gellers, Joseph (M.), 27 X				
Applications Places System 🗱 😪				III 3 45 AH 🗳
te Edit Vew Jerrarul Teps (prip				
reiningBlocalment -16 myrel -o training -p < /mome/trainin	\$/1113A5.091			
ter passwort				
rachingElocathest -16 myset -o training -p ter personnet				
icome to the Hybbi monitor. Commands and with ; or bp.				
or MySQL connection id is 5				
rver version: 5.0.77 Seurce distribution				
an inelaci or the dar bela. Type that to clear the buffer				
to state as its state that or in case or states				
ngia une meximient;				
ading table information for completion of table and column				
o can turn off this feature to get a quicker startup with	-4			
fabuse changed				
sale select - from treat limit h:				
CORTAL 1 HOLFUTAL ANN	1 CIT 1 DAT	PATIENT I PATIENT NAM AL	I GENERAL PROPERTY I TEMPERATURE R	
A BF RIN BF RA INTERNCT WITH DICTO LAB TES		REPORT I TREATMEN	MARCE PRESCRIPTION	1 821
a l'acter l'acter l'acter d'acter l'acter l'acter				
1 80 1 100 1 ORDS THE PATIENTS ACTN SUPPRESS	DENKI 8/1/2018	1243 ABH2MAINN 26	R FEIER 97	1 10
	0 040441 \$/12/3415	1 LOSS MALLANDAN 15	ET SEEDIG REPORTS DESCRIBED TO PATIENT	1 47
1001 HILAN HOSPITAL				
1005 #SIAN HOSPITAL	Overval 38/32/395	1 5564 ABH25EK8 45		1 333

SQOOP

==========================**How** To Import Table(phonedata)Mysql Database To Hadoop[HDFS]:Using Sqoop Tool _____

==

[tı	raining@localhost	~]\$sqoop	import	connect
jd	bc:mysql://localho	st/movielen	s -	-username
tra	ainingpassword	trainingta	able trea	atm 1
ta	rget-dirclinc;	//.	2	

	/ /		
			educe.ImportJobBase:
Beginning im			
			Client: Running job:
job_2018021			
18/02/14 22	:54:54 IN	FO mapred.Jo	bClient: map 0%
reduce 0%			
18/02/14 22:	55:02 INF	O mapred.Job	Client: map 100%
reduce 0%			
18/02/14 22::	55:02 INF	O mapred.Job	Client: Job complete:
job_2018021			
			lient: Counters: 12
18/02/14 22:5	55:02 INFC) mapred.JobC	lient: Job Counters
18/02/14	22:55:02	INFO	mapred.JobClient:
SLOTS_MIL	LIS MAP	S=6536	
		00000	
		O mapred.Job	Client: Launched
			Client: Launched
18/02/14 22::	55:02 INF	O mapred.Job	Client: Launched mapred.JobClient:
18/02/14 22:: map tasks=1	55:02 INF 22:55:02	O mapred.Job	
18/02/14 22:: map tasks=1 18/02/14	55:02 INF 22:55:02 LIS_REDU	O mapred.Job INFO UCES=0	
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL	55:02 INF 22:55:02 LIS_REDU 22:55:02	O mapred.Job INFO UCES=0	mapred.JobClient:
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL 18/02/14	55:02 INF 22:55:02 LIS_REDU 22:55:02 punters	O mapred.Job INFO UCES=0	mapred.JobClient:
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL 18/02/14 FileSystemCo	55:02 INF 22:55:02 LIS_REDU 22:55:02 Dunters 22:55:02	O mapred.Job INFO UCES=0 INFO INFO	mapred.JobClient: mapred.JobClient:
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL 18/02/14 FileSystemCo 18/02/14 HDFS_BYTH 18/02/14	55:02 INF 22:55:02 LIS_REDU 22:55:02 Dunters 22:55:02 ES_READ= 22:55:02	O mapred.Job UCES=0 INFO INFO =87 INFO	mapred.JobClient: mapred.JobClient:
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL 18/02/14 FileSystemCo 18/02/14 HDFS_BYTH	55:02 INF 22:55:02 LIS_REDU 22:55:02 Dunters 22:55:02 ES_READ= 22:55:02	O mapred.Job UCES=0 INFO INFO =87 INFO	mapred.JobClient: mapred.JobClient: mapred.JobClient:
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL 18/02/14 FileSystemCo 18/02/14 HDFS_BYTE 18/02/14 FILE_BYTE 18/02/14	22:55:02 INF 22:55:02 LIS_REDU 22:55:02 Dunters 22:55:02 ES_READ= 22:55:02 S_WRITTI 22:55:02	O mapred.Job UCES=0 INFO INFO =87 INFO EN=66226 INFO	mapred.JobClient: mapred.JobClient: mapred.JobClient:
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL 18/02/14 FileSystemCo 18/02/14 HDFS_BYTE 18/02/14 FILE_BYTE 18/02/14 HDFS_BYTE	55:02 INF 22:55:02 LIS_REDU 22:55:02 0unters 22:55:02 ES_READ= 22:55:02 S_WRITTI 22:55:02 ES_WRITTI	O mapred.Job INFO UCES=0 INFO =87 INFO EN=66226 INFO TEN=7972205	mapred.JobClient: mapred.JobClient: mapred.JobClient: mapred.JobClient: mapred.JobClient:
18/02/14 22:: map tasks=1 18/02/14 SLOTS_MIL 18/02/14 FileSystemCo 18/02/14 HDFS_BYTE 18/02/14 FILE_BYTE 18/02/14 HDFS_BYTE	55:02 INF 22:55:02 LIS_REDU 22:55:02 0unters 22:55:02 ES_READ= 22:55:02 S_WRITTI 22:55:02 ES_WRITTI	O mapred.Job UCES=0 INFO INFO =87 INFO EN=66226 INFO	mapred.JobClient: mapred.JobClient: mapred.JobClient: mapred.JobClient: mapred.JobClient:

www.ijcrt.org

18/02/14 22:55:02 INFO mapred.JobClient: Map input records=40002
18/02/14 22:55:02 INFO mapred.JobClient: Spilled Records=0
18/02/14 22:55:02 INFO mapred.JobClient: Map output records=40002
18/02/14 22:55:02 INFO mapred.JobClient: SPLIT_RAW_BYTES=87
18/02/14 22:55:02 INFO mapreduce.ImportJobBase: Transferred 7.6029 MB in 10.749 seconds (724.2875 KB/sec)
18/09/24 22:02:12 INFO mapreduce.ImportJobBase: Retrieved 49999 records.
Contents of directory /user/training OUTPUT
Name Type
CLINCdir
Contents of directory /user/training/CLINC/part- m00000
Name Type
_SUCCESS file _logs dir part-m-00000 file
File: /user/training/CLINC/part-m00000

Edt Yew Higtory B					and over all
	okriata (Soots Likelip			
🔶 + 😢 🖸 🄮	1 10p	/locathost 5007	l/iFshealth	(NP	
adoop NarveNode localh	est.loc 4				
meNode 'loc			main:	8020	,
relan: 0.33.2-cm mpiled: The Oct 1.3	123 37 31 P	DY 2017 e 600% (ratefield 2011) by root fro	1373+79-00-4	(67)044	
a the filesystem note Laga					
r Summary 11 and directories, 134 Ngured Capacity Veed	blacks - 368	total. Heap Size	10 T	. / 148.)	# H# (2%)
on DFS Used		5.661	18		
F5 Remaining F5 Veed%		0.39			
5 Remaining%		66.97	S		
e Nodes ad Hodes		1	<u>1</u>		
icommissioning Noder			0		
her of Under-Replic	sted Blocks		49		
ede Storage:					
age Directory		Type		State	
wAbAadoop-0.30rcacheA	adooprittuis	IN MAGE AN	D_EDITS	Active	
wAbAadrop-0 30xacheA	adoopriffuns	MAGE,AN	0,6DITS	Active	
w Albihadoop O 30rracheA M	adoop//fis/w	MAGE,AN	0,60/75	Active	
	edoopriffune		D_EDITS		ill (trainin
					il (ranin
			tadoop Nar	reft]	
i taining	🤉 nek (=)		tadoop Nar	reft]	ili (trainin ing : Retil
training ee Hybry (pokmarks	🤉 nek (=)		tadoop Nar	reft]	
training (ex Hybry (polimatic) rank (4		tadoop Nar	nett]	org - Mart
training) ieu Hybry (poimato Rý) 🚯 🕅 и) rank (4	lesito 🛯 👹 s	tadoop Nar	nett]	org - Mart
training (ee Hypry (pokinalis • Rg () 😭 🗗 u anhanny	() rank () d () dra pro- p dicator to #	lesito 🛯 👹 s	tadoop Nar	nett]	org - Mart
B training	() rank () d () dra pro- p dicator to #	lesito 🛯 👹 s	tadoop Nar	nett]	org - Mart
 training training the Hypey Boetnaks the Good Control of the Hypey Boetnaks the Hypey Boetnaks the Hypey Boetnaks 	(2 rank (of 3min pro publication to #	lesito 🛯 👹 s	tadoop Nar	nett]	org - Mart
e Towning At your Highey Bookmarks O + All () At Towning Its Asserburing sits of directory source/cables	(2 rank (of 3min pro publication to #	lesito 🛯 👹 s	tadoop Nar	nett]	org - Mart
training training training training training training training training training	() rank () Join pro p. docarout to #	iteshto	tadoop Nar Kor/La	vett - 1	ing Hell
tore Highly Bollmach to the Highly Boll Highly Boll to the Highly Bo	() rank () Join pro p. docarout to #	Jestino	tadiog Nar Kor), exclusion	tuner	6rep
	() rank () Join pro p. docarout to #	iesine	Hadoop Nar Loots Handleschry Pensisales	ration of the second se	6000 Filter
the second	() rank () Join pro p. docarout to #	Headford Statistics	Hardoop Nar Looda Harabalas Anna - A Marabalas	velt	free for the second sec
tearning	() rank () Join pro p. docarout to #	Headfording Time (() () () () () () () () () (100000 Nor 1000000	energy () () () () () () () () () () () () () () (67949 19979178
transmy ()	() rank () Join pro p. docarout to #	Andream 100 754m Andream 100 754m Andream 100 754m Andream 100 754m Andream 100 754m Andream 100 754m	Hardoop Nar Looda Harabalas Anna - A Marabalas	energy () () () () () () () () () () () () () () (free for the second sec
The Harring The Harring The Harring The Harring The Harring The Annual An	() rank () Join pro p. docarout to #	Headfording Time (() () () () () () () () () (100000 Nor 1000000	energy () () () () () () () () () () () () () () (67949 19979178
Transg	() rank () Join pro p. docarout to #	Andream 100 754m Andream 100 754m Andream 100 754m Andream 100 754m Andream 100 754m Andream 100 754m	40000 Nor 10000 Nor 10000 Nor 10000 Nor 1000 Nor	1997 (67449 1944 1947 1947 1947 1949 1947 1949 1947 1949 1947 1949 1947 1949
training training the Highery Bookinetic	() rank () Join pro p. docarout to #	Headdon	Panalaalaa Hoora a Hoora a Hor	1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444)	free Processon free P
tearray	(2 rank (-))	Headdon		1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444)	
per lapor guarante t actury constraints t d'actury constraints actury constraints	(2 rank (-))	Headdon		1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444) 1444)	

FUTURE ENHANCEMENTS

Apache Spark is an open source processing engine built around speed, case of use, and analytics. If you have large amounts of data that requires low latency processing that a typical Map Reduce program cannot provide, Spark is the alternative. Spark provides in-memory cluster computing for lightning fast speed and supports Java, Scala, and Python APIs for ease of development.

CONCLUSION:

In this paper, we presented a study on Clincal is help to doctors by getting idea what treatment he has given to that particular patient. To analysis theClinical data in hadoop ecosystem.Hadoop ecosystem is hive, pig, mapreduce. In future the spark 100 times faster than hadoop, it is easily analysis faster.

REFERENCES:

[1] Hermann M, Pentek T, Otto B. Design Principles for Industrie 4.0 Scenarios, pp. 3928-3937, 2016.

[2] Thuemmler C, Bai C. Health 4.0: How Virtualization and Big Data are Revolutionizing Healthcare, Springer International Publishing, 2017.

[3] Cao P, Xian Medical University. "The Impact of Big Data on the Development of Chinese Health Service Industry", Journal of Electronic Test, 2014.

[4] M. Akay, D. I. Fotiadis, K. S. Nikita and R.
W. Williams, "Guest Editorial: Biomedical Informatics in Clinical Environments," Biomedical and Health Informatics, IEEE Journal of, vol. 19, no. 1, pp. 149-150, 2015.

[5] Black A D, Car J, Pagliari C, et al. "The impact of eHealth on the quality and safety of health care: a systematic overview", PLoS Medicine, vol. 8, no.1, 2011.

[6] Horsky J, Schiff G D, Johnston D, et al. "Methodological Review: Interface design principles for usable decision support: A targeted review of best practices for clinical prescribing interventions", Journal of Biomedical Informatics, vol. 45, no. 6, pp. 1202-1216, 2012.

[7] Valdez A C, Ziefle M, Verbert K, et al. "Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives", Machine Learning for Health Informatics, Springer International Publishing, 2016.

[8] Yu W D, Pratiksha C, Swati S, et al. "A Modeling Approach to Big Data Based Recommendation Engine in Modern Health Care Environment", Computer Software and Applications Conference, IEEE Computer Society, pp.75-86, 2015.

[9] Bhatti U A, Huang M, Zhang Y, et al. Research on the Smartphone Based eHealth Systems for Strengthing Healthcare Organization[M]// Smart Health. 2017.

[10] Vancampfort D, Mugisha J, Hallgren M, et al. "The prevalence of diabetes mellitus type 2 in people with alcohol use disorders: a systematic review and large scale meta-analysis", Psychiatry Research, vol. 246, pp. 394-400, 2016.

