SRMS College of Engineering & Technology, Bareilly

Vol. 9
Issue-9

September-2021

E-NEWS LETTER

Master of Computer Applications

Contents:

LATEST UPDATES WORLD

Next Generation Reservoir Computing: A New Way To Solve the "Hardest of the Hard" Computer Problems

Artificial neural networks – the heart of reservoir computing – have been greatly simplified.

A relatively new type of computing that mimics the way the human brain works was already transforming how scientists could tackle some of the most difficult information processing problems.

Now, researchers have found a way to make what is called reservoir computing work between 33 and a million times faster, with significantly fewer computing resources and less data input needed.

In fact, in one test of this next-generation reservoir computing, researchers solved a complex computing problem in less than a second on a desktop computer.

Using the now current state-of-the-art technology, the same problem requires a supercomputer to solve and still takes much longer, said Daniel Gauthier, lead author of the study and professor of physics at The Ohio State University.

“We can perform very complex information processing tasks in a fraction of the time using much less computer resources compared to what reservoir computing can currently do,” Gauthier said.

“And reservoir computing was already a significant improvement on what was previously possible.”

The study was published on September 21, 2021, in the journal Nature Communications.
Reservoir computing is a machine learning algorithm developed in the early 2000s and used to solve the “hardest of the hard” computing problems, such as forecasting the evolution of dynamical systems that change over time, Gauthier said.

Dynamical systems, like the weather, are difficult to predict because just one small change in one condition can have massive effects down the line, he said.

One famous example is the “butterfly effect,” in which – in one metaphorical example – changes created by a butterfly flapping its wings can eventually influence the weather weeks later.

Previous research has shown that reservoir computing is well-suited for learning dynamical systems and can provide accurate forecasts about how they will behave in the future, Gauthier said.

It does that through the use of an artificial neural network, somewhat like a human brain. Scientists feed data on a dynamical network into a “reservoir” of randomly connected artificial neurons in a network. The network produces useful output that the scientists can interpret and feed back into the network, building a more and more accurate forecast of how the system will evolve in the future.

The larger and more complex the system and the more accurate that the scientists want the forecast to be, the bigger the network of artificial neurons has to be and the more computing resources and time that are needed to complete the task.

One issue has been that the reservoir of artificial neurons is a “black box,” Gauthier said, and scientists have not known exactly what goes on inside of it – they only know it works.

The artificial neural networks at the heart of reservoir computing are built on mathematics, Gauthier explained.

“We had mathematicians look at these networks and ask, ‘To what extent are all these pieces in the machinery really needed?’” he said.
In this study, Gauthier and his colleagues investigated that question and found that the whole reservoir computing system could be greatly simplified, dramatically reducing the need for computing resources and saving significant time.

They tested their concept on a forecasting task involving a weather system developed by Edward Lorenz, whose work led to our understanding of the butterfly effect.

Their next-generation reservoir computing was a clear winner over today’s state—of-the-art on this Lorenz forecasting task. In one relatively simple simulation done on a desktop computer, the new system was 33 to 163 times faster than the current model.

But when the aim was for great accuracy in the forecast, the next-generation reservoir computing was about 1 million times faster. And the new-

generation computing achieved the same accuracy with the equivalent of just 28 neurons, compared to the 4,000 needed by the current-generation model, Gauthier said.

An important reason for the speed-up is that the "brain" behind this next generation of reservoir computing needs a lot less warmup and training compared to the current generation to produce the same results.

Warmup is training data that needs to be added as input into the reservoir computer to prepare it for its actual task.

"For our next-generation reservoir computing, there is almost no warming time needed," Gauthier said.

"Currently, scientists have to put in 1,000 or 10,000 data points or more to warm it up. And that's all data that is lost, that is not needed for the actual work. We only have to put in one or two or three data points," he said.

And once researchers are ready to train the reservoir computer to make the forecast, again, a lot less data is needed in the next-generation system.

In their test of the Lorenz forecasting task, the researchers could get the same results using 400 data points as the current generation produced using 5,000 data points or more, depending on the accuracy desired.

"What's exciting is that this next generation of reservoir computing takes what was already very good and makes it significantly more efficient," Gauthier said.

He and his colleagues plan to extend this work to tackle even more difficult computing problems, such as forecasting fluid dynamics.

"That's an incredibly challenging problem to solve. We want to see if we can speed up the process of solving that problem using our simplified model of reservoir computing."

DEPARTMENT ACTIVITY

1. Zero hour activity on 23rd September 2021
(Departmental Poster Making Competition):

Venue: MCA Class Room

MCA department organized Poster Making Competition in the zero hour activities in which students were divided into three groups A, B & C.

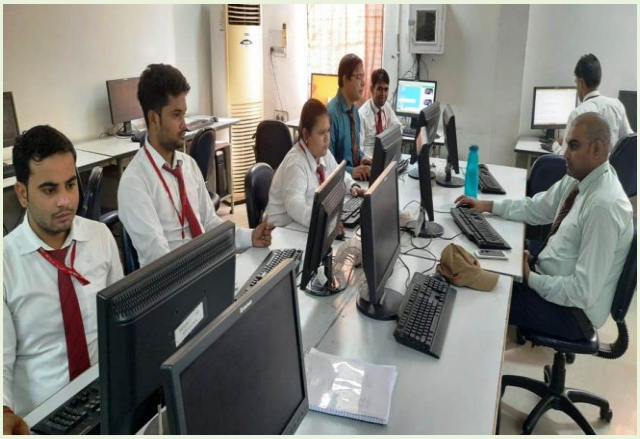
Drawing sheet had provided by the department on which they had to draw their ideas through poster on the given topic "Global Worming".

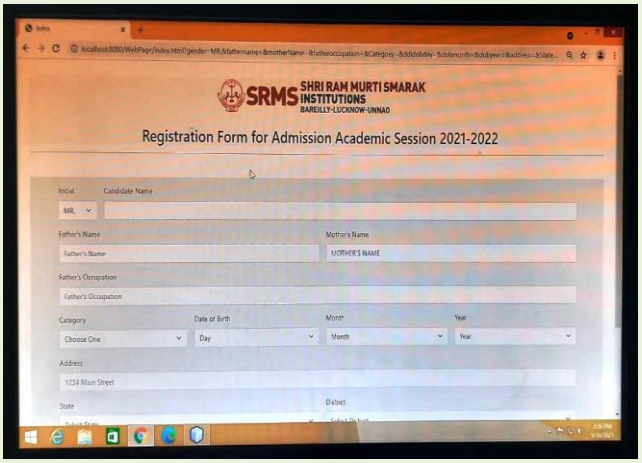
Group-A (Arsh Rupainwar & Sachin Saxena) emerged out as a winner followed by the second and third group. This event was conducted under the supervision of MCA faculties Dr. Sanjay Kumar Pandey, Mr Arvind Kumar Mishra and Mr Vijay Kumar Dubey.

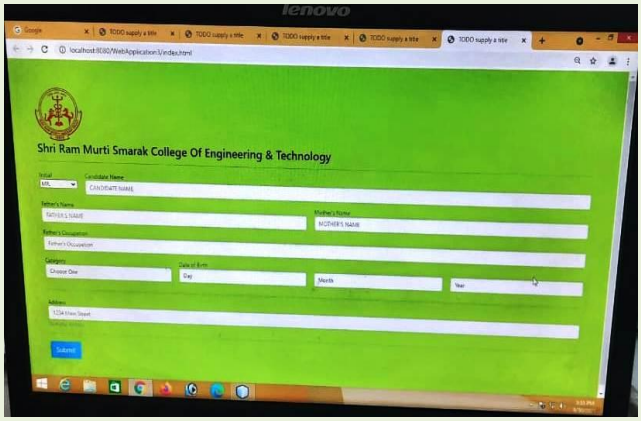
2. Zero hour activity on 30th September 2021
(Web development training and competition)

Venue: MCA Computer Lab-B

On 30th February, 2021 MCA department organized a web development training and competition in the club activity hour (zero hour), in which Dr. Sanjay Kumar Pandey, Mr. Arvind Kumar Mishra & Mr. Vijay Kumar Dubey introduced students about web development and trained them how to develop a webpage. Then they organized a competition between students to make their own attractive webpage individually on editor and host the page on **000webhost.com.**





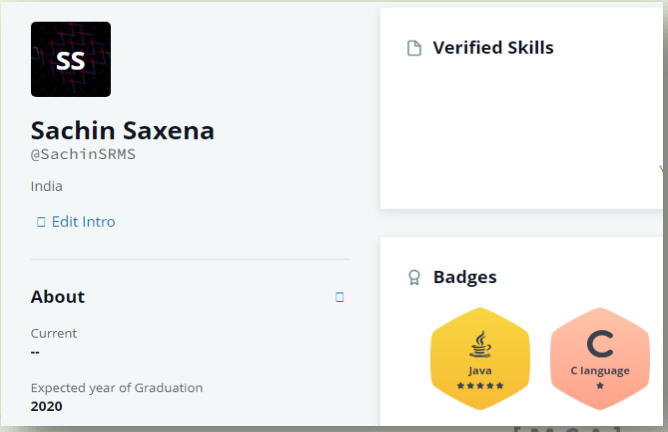


In this competition **Jaswant Singh** got the first position by making the best webpage and second position is occupied by **Gaurav Mishra.**

STUDENT CORNER

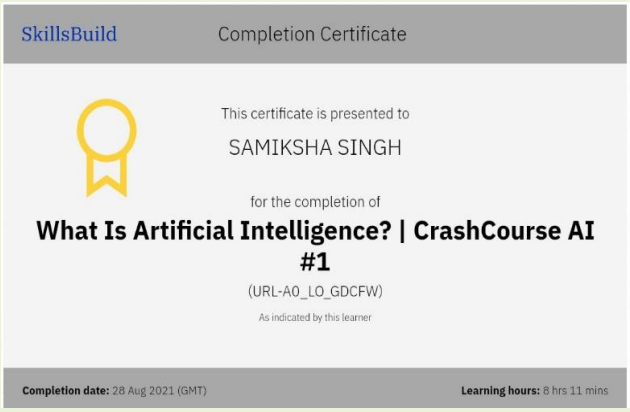
1. Sachin Saxena, student of MCA 2nd year achieved 5-stars (*****) in Java Language proficiency from hacker rank.

2. Ashutosh Maurya, student of MCA 2nd year



achieved Cloud Practitioner Essentials certificate from AWS.
AWS is the updated digital course is for individuals who want to develop a fundamental understanding of the AWS Cloud, independent of any specific technical role. You'll learn about AWS Cloud concepts, core AWS services, security, architecture, pricing, and support to build your AWS Cloud knowledge. This course will also help you prepare for the AWS Certified Cloud Practitioner exam.

3. Samiksha Singh, student of MCA 2nd year achieved Crash Course AI certificate from IBM Skills Build.


SkillsBuild Completion Certificate

This certificate is presented to

SAMIKSHA SINGH

for the completion of

What Is Artificial Intelligence? | CrashCourse AI #1

(URL-A0_LO_GDCFW)

As indicated by this learner

Completion date: 28 Aug 2021 (GMT) **Learning hours:** 8 hrs 11 mins

FACULTY CORNER

PolyBase in SQL Server

One of the more intriguing features included with SQL Server is PolyBase, a transparent access layer that facilitates connectivity between the database engine and external data sources containing unstructured or semi-structured data. PolyBase is optimized for data warehouse workloads and analytical query processing, making it easier than ever to merge big data into the SQL Server universe. It is a technology that accesses and combines both non-relational and relational data, all from within SQL Server.

If you want to learn how to use SQL Server, from beginner to advanced techniques, then visit Mindmajix- A Global online training platform: "SQL Server Training" Course. This course will help you to achieve excellence in this domain!

What is PolyBase

PolyBase was developed at the Microsoft Jim Gray Systems Lab at the University of Wisconsin-Madison under the direction of Dr. David DeWitt, a Microsoft Technical Fellow. It provides an interface that allows you to work with data stored in HDFS by using SQL syntax in PDW queries—in a manner similar to querying a linked server from SQL Server rather than MapReduce jobs. You can even use PolyBase to join relational data in PDW with data in HDFS, as shown in figure bellow. In addition, you can use PolyBase to move data from PDW to HDFS or vice versa. Furthermore, you can use Power Query or Power Pivot to connect to PDW and use PolyBase to import data from HDFS into Excel.

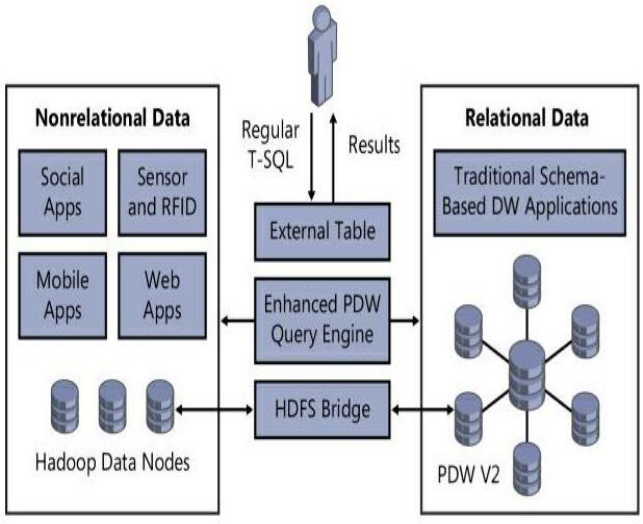


Figure: PolyBase as an HDFS bridge between Hadoop and PDW.

Benefits of PolyBase

The most obvious benefit of the availability of PolyBase in PDW is the ability to combine both relational and nonrelational data into a single

result set, but there are several others. In particular, database professionals already familiar with developing SQL queries to retrieve data from PDW for reporting and analytical applications have nothing new to learn when they need to query non-relational data. There is no need to learn MapReduce, nor is there any need to learn how to use the other tools in the Hadoop ecosystem, such as HiveQL, Pig, or Sqoop. Existing SQL skills are sufficient.

Another benefit is faster results from queries to HDFS. PolyBase is able to perform read and write operations in parallel much faster by taking advantage of the massively parallel processing (MPP) of PDW. Whereas using Sqoop is effective for moving data in and out of a relational database, it processes data serially and interfaces with the PDW control node. By contrast, PolyBase not only parallelizes data transfers but also moves data directly from Hadoop data nodes to PDW compute nodes, as shown in Figure bellow.

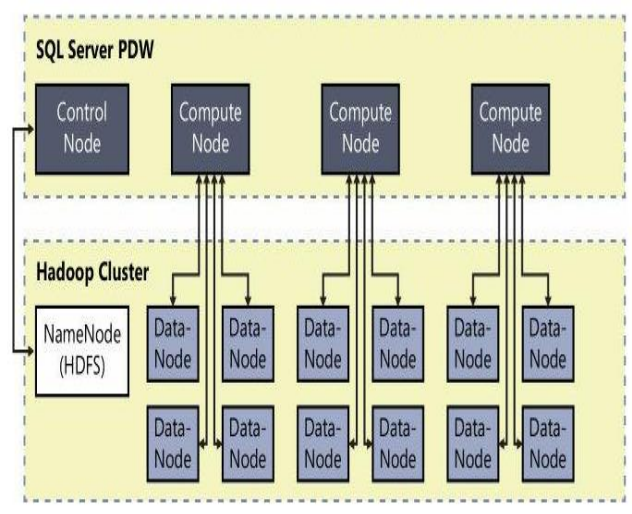


Figure: Parallel data transfer between PDW compute nodes and Hadoop data nodes.

Lastly, PolyBase is highly flexible. It is not limited to a single operating system or Hadoop distribution. In addition, it supports any type of HDFS file format. This means that you can use PolyBase to deliver data from all types of unstructured sources across the entire Microsoft BI stack. You can connect to PDW with ad hoc analysis tools like Excel and Power BI or distribute standard reports by using Reporting Services. You can even use data from PDW in Analysis Services multidimensional or tabular models to enrich it with business logic and other BI features.

Accessing HDFS data from PDW

The work that PolyBase performs to retrieve data from HDFS is transparent during queries. The only requirement is that you create an external table to define the schema that PDW can then query. You can then interact with data in HDFS files in multiple ways, either by moving data between systems or by querying both systems and joining the results.

Let's say you want to export data from a PDW table called FactInventory and store the results in a text file on your on-premises Hadoop cluster. You use Create Table As Select syntax to create an external table and transfer data from PDW into a file in HDFS, like this:

You can then write queries that reference both PDW and external tables pointing to HDFS, as shown here:

SELECT sl.machineName, m.machineDescription, m.machineStartDate, sl.eventDat FROM ServerLogs sl

JOIN DimMachine m

ON sl.machineName = m.machineName

Mr. Vijay Kumar Dubey
Assistant Professor
Department of MCA

COLLEGE UPDATE

- Regular classes begin from 14th September for the academic session 2021-22 (odd semesters).

- **Hindi Debate Competition**
  27th September, 2021
  Venue: Shri Ram Murti Smarak Centennial Auditorium, Bareilly

  The minute-to-minute programme of hindi debate competition was emceed by Shri Ram Murti Smarak College of Engineering & Technology, Bareilly. The session was carried on by the esteemed presence of Dr. Prabhakar Gupta Sir, Chairman Sir Shri Dev Murti, Aditya Murti Sir. The contingent was initiated by the melodious chants of SaraswatiVandana with Lamp Lighting glorifying the aura of the event. Sapling presentation was relayed on depicting the warmth of traditions. Chairman Sir inscribed the onlookers with his welcoming locutions. Gists of regulations were decreed and Senior Wing Debate was prosecuted.

  The debate was triumphed as:

- **For The Motion :**
  1st Shariq Shabab Zaidi (Bareilly College)
  2nd Shruti Saxena (SRMSCET Pharmacy, Bareilly)
  3rd Gaurarang Agarwal (SRMSCET, Bareilly)

- **Against The Motion :**
  1st Yatharth Anand (Bareilly College)
  2nd Kiran (Maharaja Agrasen Mahavidyalaya, Bareilly)
  3rd Anjaney Bajpayee (SRMSCET, Bareilly)

  Finally, a vote of thanks was designated to spectators by Ms. Shivangi Shukla (Tyro President) and the event culminated by deferential chorus of National Anthem.