Evaluation of teaching effectiveness in data analysis courses using a behavioural big data model

Yixu Wang

Evaluation of teaching effectiveness in data analysis courses using a behavioural big data model

Yixu Wang

Teacher Digital Literacy Enhancement Center,
Jilin Provincial Institute of Education,
Changchun, Jilin, 130000, China
Email: 15164347351@163.com

Abstract: This study proposes a behavioural big data model for evaluating teaching effectiveness in data analysis courses across primary, secondary, and higher education levels. The framework integrates learning management system data, classroom engagement indicators, and student interaction behaviours to provide a comprehensive understanding of how teaching strategies influence learning outcomes. At the primary level, the model captures early learning patterns such as task attention, problem-solving attempts, and basic data reasoning through gamified digital activities. For secondary students, behavioural indicators – including learning persistence, collaboration, and response accuracy –are used to assess the development of analytical thinking and computational skills. Machine learning and statistical techniques, such as clustering, regression, and correlation analysis, identify patterns linking teaching approaches with student performance and motivation. With a predictive accuracy of 89%, the model demonstrates strong adaptability across age groups. Findings show that interactive, feedback-rich, and project-based learning environments significantly enhance students' comprehension and retention of data analysis concepts.

Keywords: data analysis education; behavioural big data; learning analytics; teaching effectiveness; student performance evaluation; big data in education; data-driven pedagogy.

Biographical notes: Yixu Wang is a faculty member at the Teacher Digital Literacy Enhancement Center, Jilin Provincial Institute of Education, Changchun, Jilin, China. Her professional work focuses on teacher digital literacy development, educational technology integration, and the enhancement of teaching competencies through digital transformation. With experience in teacher training and educational research, she is actively involved in promoting innovative pedagogical practices and supporting educators' professional development in the context of information and communication technologies in education.

1 Introduction

Currently, technology has grown so big that it changes everything, even school. Some good changes have come about in how teachers and students are taught because of these tech changes (Jiménez et al., 2019). As we rebuild, some of the most important people who work in education are changing what they do for a living and how they do it. Tech will be used to help them learn in new ways as well. For this latest view to work, you and your kids will need to know how to use technology. Things change all the time, but kids today are more used to technology because that is how they were raised. Tech is used in school to meet the wants of kids who have grown up with it and to make them feel better (Aznar et al., 2019). And this is the main reason why it is important to teach in a way that fits the way things are done in school now. Information and communication technologies (ICTs) should be used to prepare everything from now on. It is important for teachers and students to be able to use many types of computer tools and files. It is a lot like the rules that schools have now. They are meant to make schools better. Because so many people use ICT, I can offer new ways to learn that put the student in charge of change and keeping information safe. Therefore, the teacher cannot just give out information; they have to come up with unique ways for each student to learn better and in a way that works for them (López-Belmonte et al., 2019).

It has been mentioned by many teachers and experts in the field that tech in the classroom can help kids learn. They found that technology in the school makes learning more fun, more interesting, more useful and more creative. These are the reasons why schools should use new tech. China was chosen because it is a great place to learn. From 2011 to 2020, World Development Indicators show that China spent about 21.90% more on R&D than the USA did. Their creativity score went up by 14.87% during that time. People in China like new ideas and different ways to run their businesses, so we should learn more about them. The way people think about technology and how new business models are made changes when they look at 'big data'. That is the point of our work. Some of these studies look at how workers' creativity and interest in technology affect the connection between the two. The view that changes powers work with this setup (Asemokha et al., 2019). As part of our study, we polled middle and top-level managers at Chinese companies. The information was studied using SEM, which stands for 'structural equation modelling'. Businesses that know how to look at many data can make more money and make their tech better.

Being tech-savvy and artistic at the same time can help you use big data to come up with new ideas. As people work, they come up with new ideas. Big data skills and new ideas do not mix too much when you are artistic. Many new ideas have come from this study (Breier et al., 2021). It confirmed what was already known about how companies can benefit from being able to change their rules. Many people have asked us to find out more about the forces inside companies that allow them to run in new ways. No other research has looked at how knowing how to analyse large amounts of data can help people make new business ideas. This study looks at things that hurt and things that help. Two of these reasons have to do with workers' creativity and technical skills. It is also more important to think about 'dynamic capacities' now than it was in the past. In China, no one has ever looked at that before. This is another way that studying the past has made it more interesting. There have also been big changes in how things are run because of this study. They show that in order to come up with new business models, managers need

to agree with and back up the way experts think. They should also hire people who know how to work with big data and teach them (Cui et al., 2022).

To help them grow, businesses should look for people who can think outside the box. Things will get worse when they are in charge of other people. In this case, the most important people could make a big difference. The world is changing very quickly because of all the new, cutting-edge technology out there. These days, people use a lot more tech tools, and these tools are always generating a lot of data. Some people say these apps and tools are good for saving and seeing data. A lot of people use 'big data' these days in business, education, health studies, statistics and other areas (Hadwer et al., 2019). The reason it is getting so much attention is this. 'Big data' is made up of three main ideas, which are sometimes called the 'three Vs'. These three things are number, speed and range. Cell phones, social media sites, business deals, and personal travel habits are just some of the places where a lot of data is coming from. This data is growing very quickly and all over the place. This is the case because the data is growing so rapidly and in so many areas. There are now more than three Vs in this idea of the three Vs. We have talked about how to group together big pieces of information. People sometimes call these five parts the '5Vs' (Aslam and Khan, 2020). The parts of this method are number, speed, variety, being honest and value.

The system does think about these things. They talked in great depth about 'big data' in their piece. The seven special things about it are often called the '7V's'. People are thinking about three things: volume means the huge amount of data that is created. How fast data is made and moved around is called its velocity. It also refers to the different kinds of material and sources. An important way to look at statistics is to divide them into four groups: worth, feeling, honesty and variety. Your 'valence' is how you feel when you know something, good or bad. 'Veracity' in this case means how true and reliable the story is. This is what 'variability' means: a collection, on the other hand, changes over time and has gaps in it (Chowdhury et al., 2023). Says that material has 'value' if it is useful and important. Many areas need big data a lot more now than they did in the past. These include construction, insurance, healthcare, communications and e-commerce.

The integration of data analysis education in both primary and secondary schools is essential for nurturing analytical and computational thinking from an early age. This study develops a behavioural big data model to assess teaching effectiveness by analysing learning patterns, engagement levels, and performance metrics across educational stages.

This paper's structure is setup as follows: the relevant work on the assessment of teaching effectiveness is presented in Section 2. The data analysis course-based technique is described in Section 3. Results pertaining to the big data model are discussed in Section 4. Lastly, the findings are given in Section 5.

1.1   *Contribution of the study*

With this study, a behavioural big data model for teaching data analysis is put forward and tested. It will add to the growing body of research on big data in education. Deep learning, learning analytics, and descriptive statistics are all added to a system that already works well. This makes the field of study move forward. These rules help us figure out how our students learn and how their actions affect how well we teach. The study explains how to arrange a lot of behavioural data in a way that gives useful information. This information can be used to create better lesson plans and personal learning tools. A lot of different tools, like ETL methods, comparison analysis, and

finding outliers, are used to make this possible. The study makes a big contribution to the field of methods and also helps schools and teachers a lot. It is possible to guess how well students will do in school, find students who are at risk, and help those students get the help they need to do better by using learning management systems (LMSs) and behavioural data. When kids use big data analytics in data analysis class, they learn more and better. This fits with a bigger trend toward schooling that is based on technology and facts. People who work in education can learn from these results how to use big data to create flexible, effective, and student-centred ways to teach.

2   Related work

2.1   Technology tools for teaching analysis

You should write down how you teach so that you can do 'teaching reflection', which means you think about how you teach. This is a very important thing to think about because we need to change professional teacher education, make training more fair, and get teachers to do a better job at their jobs. There is a new idea that learning analytics can help teachers learn. Teachers can use it to watch other teachers teach online (Li et al., 2013). People know a lot of information, and this field of study tries to make sense of it all. It tells teachers how well their kids are doing in school and guesses how well they will do in the future. It also finds problems that might happen. The use of learning analytics helps teachers better understand how they teach. They can tailor their lessons to each student better with the information they gather. They can use the tools for learning analytics to do this. When you learn online, the site that makes it possible saves detailed records of what students do, which is like seeing pictures of teachers at work (Wei, 2013). It is possible to copy how people learn online by looking at facts about how teachers and students act. This tells teachers everything they need to know about how to teach online, like the different parts of teaching, the jobs that are important for both teaching and learning, and details about each student.

The fuzzy picture from the beginning will be scanned and made clear. Teachers will then have to think about the good and bad things about how they teach and learn, as well as how they create tools to help students learn and keep track of their progress. A lesson on Moodle was picked to show people what to do (Han et al., 2015). The main goal of online teaching feedback is for teachers to keep track of how their students are learning from their own point of view. This has been looked into. As an online teacher, you might find it helpful to learn about SQL SERVER 2005 Analysis Service (SSAS), SPSS, UCINET, Excel, and ICTCLAS, which is a method for separating Chinese words. One can use these to see records of the lessons from the teacher's point of view. Text mining, link mining, grouping, and social network analysis are some of these tools. Data mining methods are used for statistical analysis and image analysis. This could give us a whole new way to think about how to teach and learn. It is being used for more things now that there is a new study on it. Check out how big data is used in internet star marketing to change how people choose what to buy. It can be used to study how people act, so pay close attention to that. The supply chain and business data can be safer with the help of big data. This shows that it can help with some things, like figuring out risks and managing processes.

2.2   *Predictive modelling techniques in educational data mining*

You can also guess what people all over the world will want to buy by using big data from social media sites. You can use these sites to improve your business and product ideas after reading this. Large amounts of data could have a big impact on how important business decisions are made and how well people work in many areas (Abdul Bujang et al., 2021). People in the school world have already done a lot of work to try to guess how well someone will do in college. This work has made it possible for more in-depth research to follow. As early as the 2000s, researchers from other countries started to look into how to use machine learning and educational data mining to predict how well a student would do in school. In the last ten years or so, studies on this subject have grown in scope and depth. Uses students' question-and-answer data from an online learning tool to predict and model how well they will learn. The model is based on Bayesian networks. It also gives each student their own ideas for how to teach based on how well the guess worked (Cao, 2022). These are the facts we found about the background, study topics, use cases, technical methods, and growth trends of school data mining in the USA and other countries. It was also talked about what is wrong with educational data mining in Chinese schools and how it could be used.

This kind of model is called a 'matrix decomposition-based model' if it splits the student-course rating matrix into two low-rank matrices that reveal hidden information about the student and the course. After that, they are used to guess what grades the student will get in other classes. There is a way to imagine what grade a student will get in their next class by using Bayesian nets and their past grades. Other experts have also used rule-based writing methods that worked well to finish their study projects. Guess what grade the student will get on the 'dynamics' job based on their overall grade point average and the grades they got in the four classes before this one. A website for learning talked about something called a 'deep knowledge tracking model'. This model uses a deep neural network to keep track of how well students understand certain subjects and to guess how well they will do in the future based on how they answer questions. You can get an idea of how well a student will do on the final test by looking at how well they did on their homework, midterms and course projects. Based on the 'length of video viewing', the 'average number of modules studied per day', and the 'number of quizzes completed' on the catechism platform, a model was made to predict how well the students would do.

The graph neural network method was used to show how the bits of information are linked. They were able to guess what the kids would know and see how much the kids knew. Find out how well students would be able to learn many things at once using graph neural networks and multi-task learning. Based on how they liked to learn, they were put into groups. Next, a convolutional neural network was used to look at the specifics of the course and see how each student learned compared to other students in the same group. This was done to make sure that every student could finish the study. Backpropagation is a type of artificial neural network that can study and model the things that affect college students' grades and also predict what grades they will get in the end. A decision tree was used to choose, sort, and prepare the raw data of a high school score from a different country. With tools like SK learn, the decision tree method was then used to make models and study papers that could be used offline. Lastly, the decision tree method was used to guess what the students' grades would be. Some experts looked at deep learning and meta-learning as possible ways to imagine how well students would do in online

classes (Zhang et al., 2024). They made a forecast model that changes a deep neural network's design and anything that makes it unstable on its own. These tests were done with more than 500 master's students at online schools.

2.3 *Real-time big data processing for predictive analytics*

It is about the same as other models that have been used, but it works better and can be used on a larger scale. The time it takes to look at data and mine input should be cut down, and the results should be more accurate. They added Twitter data in real time so that they could get quick feedback from mood studies. This is how the computer can get the study information. One of the most important parts of data analysis is checking to see how accurate the results are. When it comes from social media, the Hadoop system gets it right more than 84% of the time. Every day, you can find out a lot about Twitter. People all over the world know about this social networking site. They could look at the info with the 'Info Sphere Big Insights' tool. Big businesses that need to look at data right away will love Hadoop's power. Data tools are a part of Bloomberg's job. Businesses benefit from this because it lets them hear from customers and find out about new trends. Airlines and a few other big names in the same field use these tools to listen to their users and figure out how to talk to them (Cui et al., 2020). It is their job to find people who have already done something wrong so they can look into it. Hadoop parts are used in a plan to find out how people are feeling on Twitter. These parts make up the Hadoop system as a whole.

So, it lets you use Hadoop computers to read tweets. It also showed a full picture of the information gathered from many people and the tweets they sent. A lot of people no longer read newspapers. Instead, they get most of their news from TV and the web. People do many other things on the web, such as buy stocks. On a PC or even a phone that can connect to the web, you can buy and sell stocks. What is going on in the stock market? They use their phones to learn about it (Jose et al., 2019). This lets them know when the market changes. People use Hadoop to see real-time data and predict what will happen in the market. A lot of data is used by people in business and in education. It checks out 1,024 GB or even 1,024 PB of info. With Hadoop and MapReduce, work is split up among many groups so that everyone can reach their goals. The idea behind this project is to use Hadoop to make stock market predictions. You can use Hive tools to create Hive tables while the data loads. Today, stock buyers do not get their news from newspapers and other old-fashioned sources. At this point, tech is so well-known these days. Thanks to the rise of cell phones, people can now buy and sell stocks on the internet right away. Besides that, buyers get news all the time on their phones. People who bought things also used them to talk about business, the market, and how to keep their homes safe.

Large amounts of this kind of data are called 'big data'. While they are in the air, the planes also send information to the offices or air stations all the time. This helps the people in charge of air traffic always know where the plane is and what its state is. We take care of everything on this list right away (Tariq et al., 2021). They give and receive a lot of information, and it tends to pile up quickly. This is because many planes talk to each other all the time. That is why this huge amount of data is called 'big data', and it is very hard to handle and manage. This study talks about this type of material and shows how to deal with it.

Recent research on educational data analytics highlights the growing role of big data in improving instructional quality and personalising learning. Prior models focused primarily on higher education, while studies in primary and secondary schooling have begun exploring learning analytics through student interaction data, attendance and assessment behaviour. Early-stage learners often exhibit distinct engagement trends, requiring models capable of capturing motivation, attention span and peer collaboration dynamics. The studies demonstrated the potential of behavioural analytics, but few frameworks have unified K-12 and university contexts within a comparative evaluation model.

3   Methodology

The next picture is a big picture of a full study method that makes it easy to combine current data analytics with the way learning tools are made. Setting study goals and finding problems in a planned way are the first steps in the process. Once this information is gathered, it is used to prepare a lesson and a way to teach. New technology is used in the method, such as ways to collect data that are then fed into advanced deep learning algorithms. When you add an LMS and big data analytics tools to them, they work even better. These cutting-edge computer methods are used to take a close look at the data after it has been collected. People can use these facts, findings, and thoughts right away. This is a fresh way to study education that helps researchers and teachers make smart choices by giving them choices based on a lot of information. There are both old and new ways of training in this system. The new ways use cutting-edge data science methods. 'Continuous revision' is a way to connect big data analytics to deep learning algorithms that does not need to be done all at once. This way of doing things is meant to keep improving how we teach and how we think about things. With ideas based on facts, this way can help the classroom run more smoothly. It allows schools a lot when they want to keep academic standards good.

3.1   Data collection

The information used in the study comes from two different kinds of sources. Transactional systems are in the first group. These keep track of both cash and information about how well kids are doing in school. It comes from polls of people who work in the school field (Villegas-Ch. et al., 2023), which is the second group of sources. A method called ETL is used to find the data. These steps can be planned with a number of different tools. The tool that was used for this job was Microsoft SQL Integration Service. Graphics-based programmes are easy for anyone to use, even if they do not know much about the subject. This tool is not only handy, but it is also useful. But it is the parts that change that give you different ways to change the data in a very specific way. This shows that the tool can do many things, not just easy ones. One of the steps used to get data and work with it can be seen in Figure 2. When you use I/O services for a transactional database, they take care of an 'OLE DB' data source. That is where the data on the left comes from. Users can make a link string to any database and choose which table to get data from with this tool. A query is used to join data from different tables or change data in certain ways, like getting rid of or changing null values. The method

receives data directly from other databases. Over the course of two semesters, information was collected from three undergraduate data analysis classes.

A multicast change is also used to send the data that comes in to two different ends in the process shown in the picture. In the shift, a conditional split is used to send the rows of data to various outputs. The info is sent to a few rows in the first item. Because of this, you can keep track of how many rows have been added to an area. The info also goes to the DataMart's levels, which are multiple data centres – a data centre with lots of different places. The ETL gets the information it needs from the Excel file on the right. This might be because the information comes from a study or other non-traditional sources and was not put together in the usual way.

Figure 1 A comprehensive research methodology framework integrating problem identification, course design, deep learning algorithms, and big data analytics to generate evidence-based educational recommendations (see online version for colours)
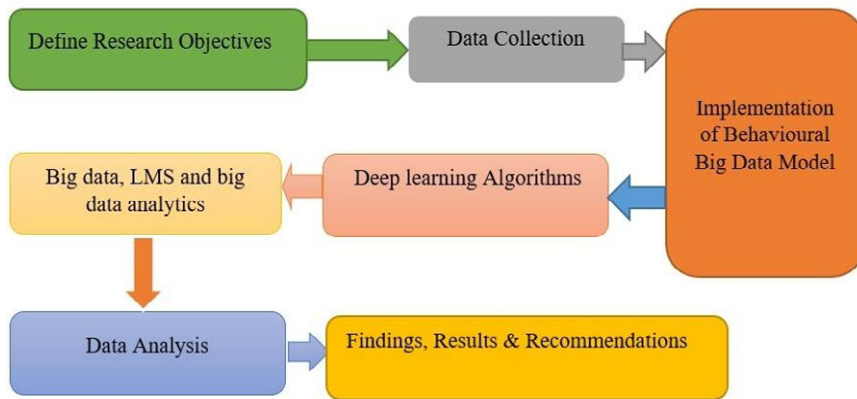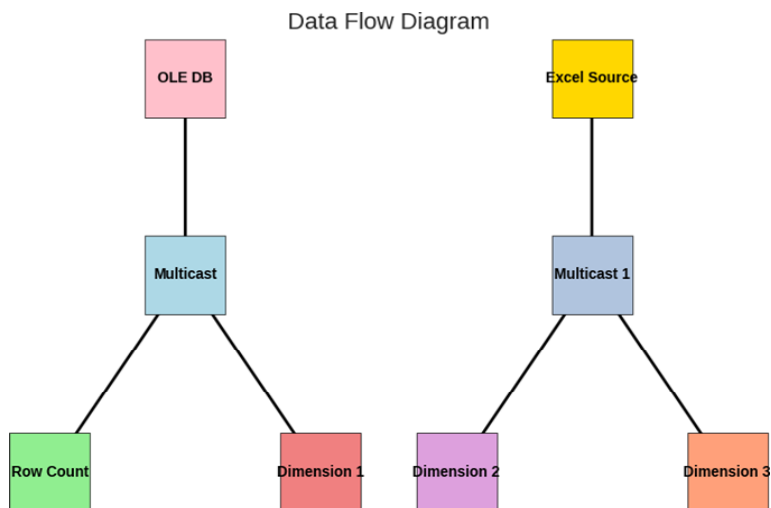

Figure 2 Extraction, transformation, and loading of data through the integration services of Microsoft SQL (see online version for colours)
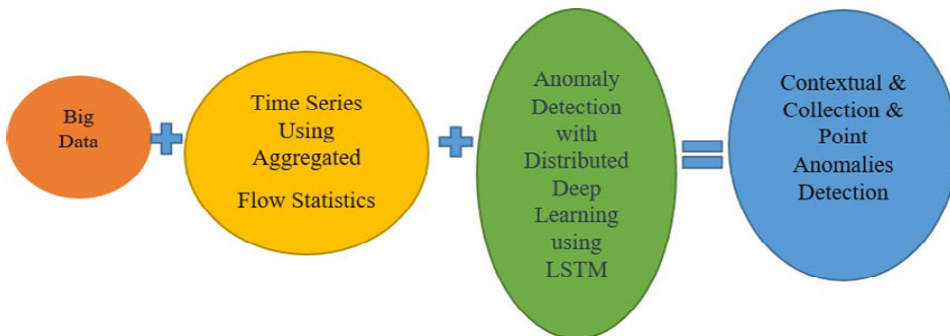
The kid's say that the polls are meant to help us figure out what really affects learning. A lot of research papers and studies that looked at what makes students learn in different ways were read and used to choose the parts. One thing that is done as part of this is a correlational study.

3.2   Implementation of behavioural big data model

Linking up, one of the deep neural networks we offer is a chatbot. It can handle both flow traffic and traffic that is put together. On top of the Apache Spark framework, long short-term memory (LSTM) is used. The Apache Spark framework helps this neural network do its job. This leads to the creation of a language with only two words: regular and odd. Along with big data, the spread of deep learning, and finding outliers in flow analysis, we want to use the ideas from language processing and add them to our plan. We are going to make a model of how the network usually works, as shown in Figure 3, so we can find outliers that are point, group, or contextual – finding patterns in large amounts of data and deep learning work really well together. This answer might be the best one. A deep learning collection needs millions of samples in order to make a big model of how people normally act. Little models that only look at outliers are not as good as this one because it often gets things wrong. Big data lets us handle this, which is what we need to do to build them (Al Jallad et al., 2019). We will be able to look at longer periods of time than we could before if we use both big data and time series. In addition, if we use this technology in IDS, we might be able to find difficult risks that have been hidden in the system for too long, maybe even months or years.

Figure 3   Proposed method (see online version for colours)

Looking at just 90 days of regular hack detection systems is not enough to find such widespread issues. This is because APT attacks take a long time to happen.

3.3   Big data, LMS and big data analytics

There are tools for 'big data' that could make it easy to store, look at and share data. Getting a lot of data changes how people learn and how they use new tools that are helpful. A picture called Figure 4 can be used to give you a general idea of what human information management's (HIM's) big data analytics is all about. A lot of people believe that LMSs like Forma LMS, Moodle, and Blackboard are great places to find a lot of information. On these sites, it is also easy to plan, carry out, keep an eye on, and grade

the learning. A lot of people use these two LMSs in the school. Manage digital materials and which students get them; manage tests and how much students learn; and manage student notes and participation. These are the three main things that this LMS programme does (Fahd and Miah, 2023).

A lot of information is made by the LMS, so we need to find better ways to teach and learn. This is also because more and more people need tools that work with LMS to read and use the data that LMS creates. LMSs are used by lots of different people for lots of different reasons. This is where big info comes from. They can use the LMS to do things like turn in their review papers or take tests in a certain area. This is what people do when they share information so that their students can use it as digital course materials, or when they use these tools to learn. You can look at huge amounts of educational data in a number of ways to find useful information and ideas. This is called 'big data analytics'. It could be about how students learn, how they log in, how long they spend on a task or module, how they turn in their tests, what page or material they read the most, when they finish a task or module, or when they post about what they do outside of school. Many pieces of information can be looked at together to get a more complete picture of how students really learn than with normal ways. Big data analytics can help HIM make better choices by using the huge amounts of data that LMS sends.

Figure 4 Overview of big data analytics in HIM (see online version for colours)
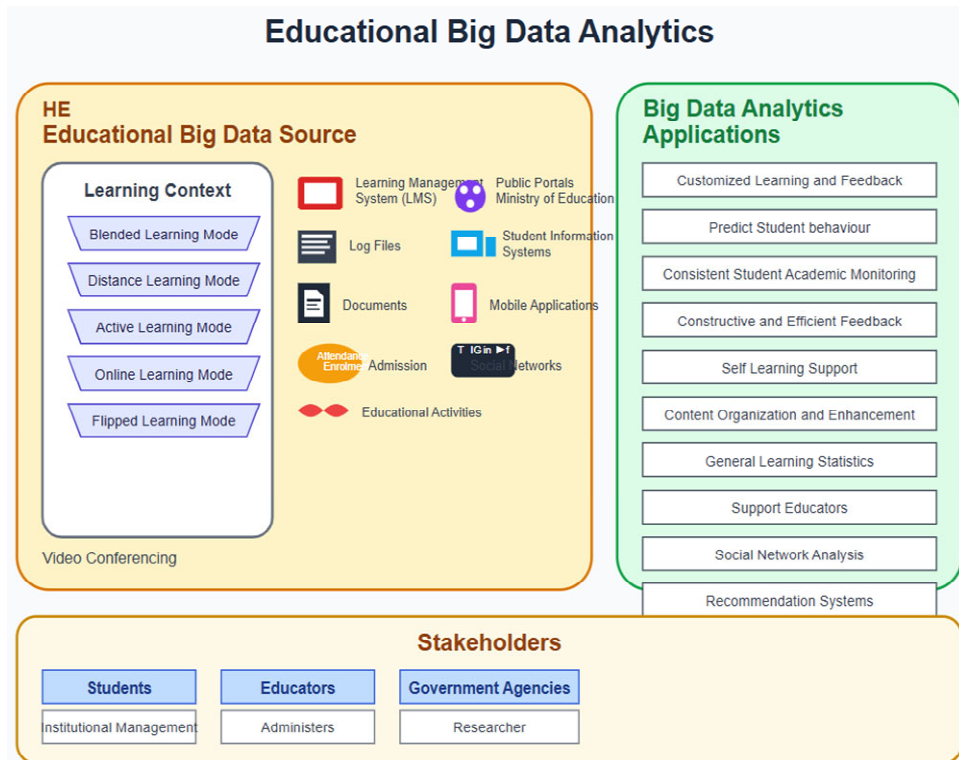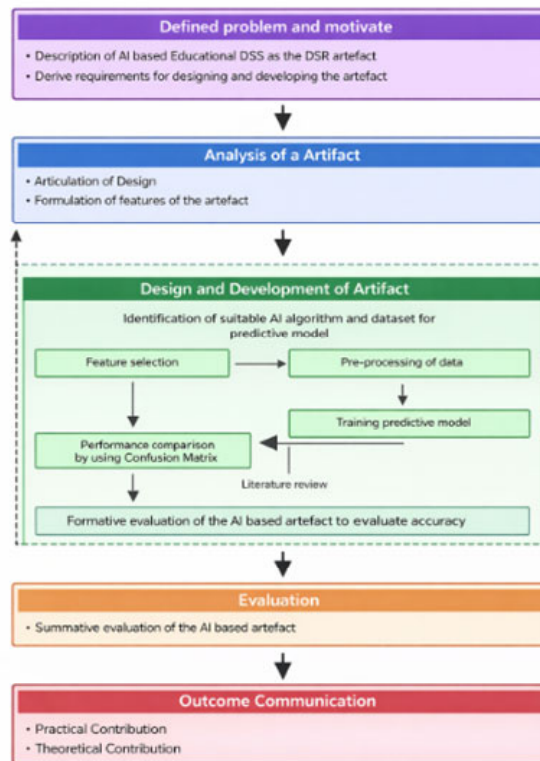
Figure 5   Integrated DSR methodology (see online version for colours)

You can use this tool to: lessons that are made just for each student to do better in school, To get kids to be more honest at school by checking their work for plagiarism; guessing how well students will do in class to figure out how to teach better; a system for picking a plan or suggesting one; finding kids who are at risk because of their behaviour so that the right steps can be taken at the right time; figure out which kids are most likely to quit; seeing how active and interested kids are in order to help them learn more; making plans ahead of time to reach college goals. Artificial intelligence (AI) gets a lot of information at once and works on it all at the same time. This is not possible in big data analytics since there is a lot of data that changes quickly. This problem can be fixed in a number of ways, and AI techniques can be used to learn from a lot of data. Using fast computers, simultaneous processing, and/or data split platforms are a few examples of this. You can do it either way. For this job, the big data analysis system (BDAS) object should be setup with one that can work with data. A way to learn that combines science and design. The research method is all about making the rules and boundaries for a study.

From a scientific point of view, this proves that the study is important and correct. They think that the method is the most important part of the study so that the goals can be met. The DBR method is used for design-based research, and the DSR method is from the field of information systems. This led to the creation of a new way to study IS that is now used. As an IT and DSR item, the goal is to learn how to make and grade a BDAS. People look at DBR as an example of how DSR can work in the education field. DSR is

an update to DBR that lets you learn the basics in new ways. One of its parts is the study of social pragmatics. A DSR method can be used for both studies that help make things and studies that give experts the information they need. They used the DSR method to make the engineering school better. One more thing they did with the DSR method was to write about how they made a mobile health information system. Someone made a phone game for learning using the DSR method. The goal of the DBR method is to help students learn more, understand how to teach and learn better, or do other things that help them learn. Most of the things that are the same both ways: you can do things in both of these ways:

- both ways of making are reasonable and doable

- both add to the method based on knowledge.

For both of these jobs, you have to think about the theory, make things that are both academic and useful, and keep going through a loop of making things and carefully reviewing them. For the study project, a mixed DSR method was used. This five-step method came about because of the things that DSR and DBR had in common. It was also used, but not in the same way as Puffer's DSR methodology. To begin, the problem is named. After that, the answer is looked over. The work is then planned and done. It has finally been looked at. In the end, the result is shared. Figure 5 shows these five steps.

In the first part of the study, the situation is fully described. Previous research is also looked at to figure out the main goal and design criteria for making a BDAS with data from the field. At this point, the systematic literature review and the meta-analysis are both finished. This helps with coming up with patterns that will be used to make the DSR product later on. After that, the outcomes are studied to come up with design ideas for the BDAS structure. This is the third step. Several AI data analysis methods are used to plan, build, and test the BDAS as a DSR product. Some of these are deep learning and machine learning. The last step is the official review, where the study data are added to what is already known.

3.4 Data analysis methods

This study will use the right ways to look at data to do the math and figure out what it all means. The study's goal is to find out how big data changes how people who are learning English are rated and how that changes how English is taught. How do you look at detailed statistics? Use descriptive statistics to figure out what all the data means after you have put it all together and looked it over. Figures such as the mean, the standard deviation, and the frequency should be found as part of this study. This research looks at how long kids learn, how interested they are in their work, how they learn, and how well they do in English. This will help us learn more about these things. Check out how the test group and the control group compare: you can use tests like the independent samples t-test or non-parametric tests to find out if the changes you have seen are statistically significant. These steps need to be taken in order to compare the trial group to the control group based on certain factors. The goal of this research is to find out how different ways of teaching English affect how well students learn to speak English (Yan et al., 2024). Based on the review of the data, both the experimental group and the control group were closely examined to determine why they learned the way they did.

We will test personalised training methods to see if they make a big difference in how long kids know, what they do while they learn, and the tools they use. We will also talk about how to make things better and what this means for those who teach.

Table 1   Data analysis method

Data analysis method	Analysis content
Descriptive statistical analysis	Learning time, learning activities, learning strategies, etc.
Comparative analysis of the experimental group and the control group	Learning time, learning activities, learning strategies, etc.
Result analysis	Explanation and discussion of the effect of the teaching intervention

The ways we have talked about above will be used to study how people learn English and how that changes how they are taught in the world of big data. You can use the results of this study to get other study results across.

A behavioural big data framework was designed to collect and analyses multi-source educational data from primary, secondary, and undergraduate data analysis courses. Using LMSs, classroom sensors, and digital platforms, features such as task completion time, consistency, collaboration and attention levels were recorded. The model combined machine learning classifiers, deep learning engagement predictors, and correlation analysis to quantify the relationship between teaching strategies and learning outcomes. Distinct behavioural indicators were weighted differently for each level – motivation and visual feedback for primary, collaboration and task depth for secondary, and performance analytics for university learners.

4   Results

A map of traits is a picture that shows what a service or product is made of. This kind of modelling is great because it lets you see quickly what the client's best qualities are. Follow your gut when you look at a picture. It will tell you which traits or features are more important and which are less important. People from studies of both kids and teachers were used to map out the traits. The project's goal is to list the most important things that get people interested in education, along with the sources that back them up. That is why we use the group of 329 people that best represents the whole community. People were asked to make a list of the seven things they think are most important at school right now. This is what the poll asked. Table 2 has them. In the first row, you can see some of the groups that worked on the project. This is the staff, and this is the group of kids. To get the numbers you know, we took the average of how important people thought each group was. It is very important to learn if the number is 10, and it is not very important if the number is 0. Let us have a chat. Some kids say it helps them learn more than what their teachers teach them. Some people believe that how well teachers use ICTs is one of the best ways to tell how well they do their job. In a mixed-methods classroom, some teachers believe that one of the most important things is that students want to learn.
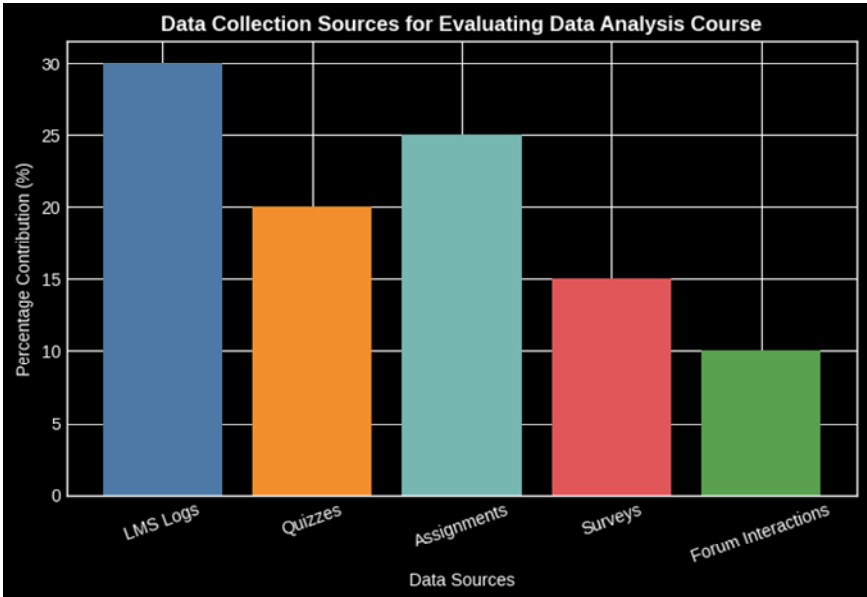
Table 2 Categories of questions for surveys of the university community and groups that participated in it

	Students group	*Teachers group*
Accessibility	9	6
Interaction	9	7
Usability	4	7
Time	9	8
Effectiveness	8	4
Motivation	8	9
Methodology	7	8

We were going to show that changes in results are made up of five parts, but the tools we had only let us show three. We worked hard for a year to try to do the study with more up-to-date tools, but we failed. We do not have any money right now because we are in Syria. Things are getting hard for us, but we still want to add to school. Why did we choose to give you this paper? We want to talk about our thoughts and how we think the tests we did went. In the same way, we hope that teachers who want to will do the whole thing over another time. The gear makes it so that tests can only be done on random parts of the collection. We will then talk about what we found and try to guess what percentage of all studies have been done so far. We will not speak about charts or numbers now because it is not the right time. The samples were all picked at random, which is why the numbers we got for each test were different. Also, the numbers are not quite right to use in charts or other types of presentations. Long-term tests of weather and group quirks are not possible because we do not have enough equipment. Different points are the only things that are checked. The chatbot model is better than the normal learning model at identifying how context will move because it has been used in more tests.

It is getting 10% less likely to give wrong answers than the old way of learning. Support vector machines (SVMs) are one of the best old-school ways to learn, so we used them to compare other models. As well as flow data, adding flow sum data to traits. Smart people add data that makes things more correct and gives a better picture of how a network works in general. This is a great pick. The fact that each second is faster than the ones before and after it could be useful. Because of what is going on around the time that is being thought about, it takes an extra second to think about it. Use of big data, deep learning, and anomaly detection is a good way to deal with overfitting, which leads to a lot of false alarms. There are new ways to find threats that are less likely to give us false results. Figure 6 shows the amount of the total score that different data sources gave to the evaluation of data analysis classes. At about 30%, LMS logs make up the largest share. Assignments come in at about 25%, and quizzes come in at about 20%. About 15% of the info comes from surveys, and only about 10% comes from group conversations. Overall, the graph shows that most of the information used to grade a course comes from system-generated behaviour data like LMS logs and tasks. At the start of our joint DSR study method, we did a full systematic literature review and meta-analysis (SLRM). AI technology in HIM was used to help kids do better in school. This was the main goal of the study.

Figure 6   Data collection framework for evaluating the effectiveness of teaching data analysis
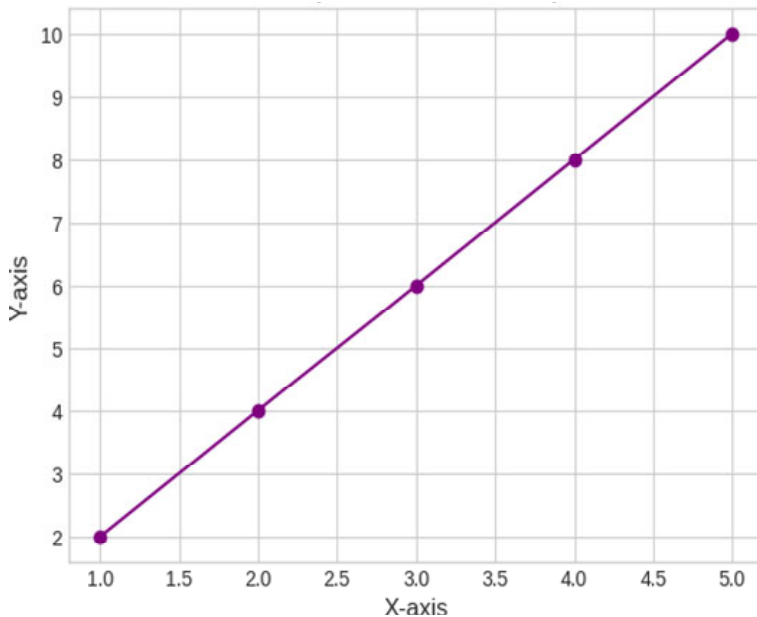course through a behavioural big data model (see online version for colours)

The thorough literature review aims to discover the most up-to-date thoughts on how to use AI-based technology to track and predict how well students will do in school. It also wants to make a list of the different AI approaches and the steps that are used to build AI models. For the SLRM, the PRISMA framework is used. A search method is laid out in this framework, along with directions on what to include and what to leave out. This helps find many things. People in the SLRM were told about the tests' steps, methods and requirements. These methods and tools for grading are very important to the start and growth of BDAS. Labelled data will help us build and improve the BDAS. It will help us train and test a prediction model that can guess how well a kid will do in school. We need to be able to find kids who are likely to fail and help them with our model for making predictions. The weather might help schools make plans that will help kids do better and learn more. You could use BDAS in your classes so that teachers can see right away how well your kids are doing in school. These tips will help kids who might be in danger a lot. Kids who are at risk can be found quickly and offered help to do better in school more quickly.

These teams can help protect kids who are in danger. For machines to learn, they need to have all of their raw information at the same time. This is because sources of big data are very big and change quickly. You cannot do this with the BDAS. This issue can be fixed with many educational data, multiple processing and fast computer technology. AI methods can then be used to build models. Many different tools can be used to split the data. There are other ways to do things as well. Because of what this study discovered, it is suggested that the BDA method design include a tool for managing and processing data. However, keep in mind that this study is not really about where the BDAS is. Instead, it is looking into how the BDAS was planned, put together and tested. Figure 4 shows one way that the DSR object was planned and made as the BDAS. The systematic literature review aims to understand the trends of application of AI-based

technology to a wide spectrum related to monitoring and predicting student academic performance and identify the different AI algorithms and process of development of AI models. A linear graph that illustrates a positive connection between the variables on the X-axis and the variables on the Y-axis is shown in Figure 7. There are five data points shown on the plot, each of which is linked by a purple line. The values range from 1 at 2 to 10 at 5. A constant increasing trend is shown by the graph, which has a slope of 2, which indicates that the value of the Y-axis grows by two units for every one unit that the value of the X-axis increases. Purple circle markers are used to indicate each data point along the linear trend, and the chart has a grid backdrop that is light grey in colour to make it easier to see the figures.

Figure 7   A simple line plot in purple (see online version for colours)

Everything that was gathered will be looked at in a broad sense through in-depth statistical analysis, which is what this study will do. In this study, descriptive statistics were used to look at things. Table 3 shows this kind of study. There are four parts to the general computer model. These are gathering data, pre-processing data, using tools to analyses data, and going over the data again. This broad model is different every time the planning and development part of BDAS is used. The BDAS used a variety of pre-processing techniques and processes to reach each step. When it comes to school big data, a lot of real-time info comes from the LMS. The study says that the BDAS forecast model is taught using old LMS data that shows how students use LMS. An open-source tool for dealing with large amounts of data is used to get new data and sort it into groups. Spark and Apache Kafka are examples of this type of software. The real-time big data pieces are sent to the BDAS so that they can be found and used to predict how well kids will do in school. When kids do this, they can do better in school and pick better things to do. You can both collect new big data and sort it into groups with a global big data processing tool. The BDAS sorts this huge amount of data into groups. In Figure 8, there

is a block-diagram process that shows how to set up a behavioural big data model. It has five steps that are done in order: collecting data, preprocessing it, building a behavioural big data model, evaluating it using metrics, and finding out how well it worked. Each block is an important step in collecting and analysing behavioural data to figure out how well teaching or learning is working.

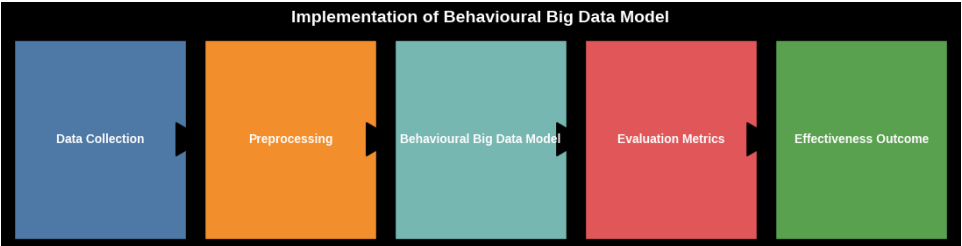**Figure 8**   Overview of BDAS as a DSR artefact (see online version for colours)
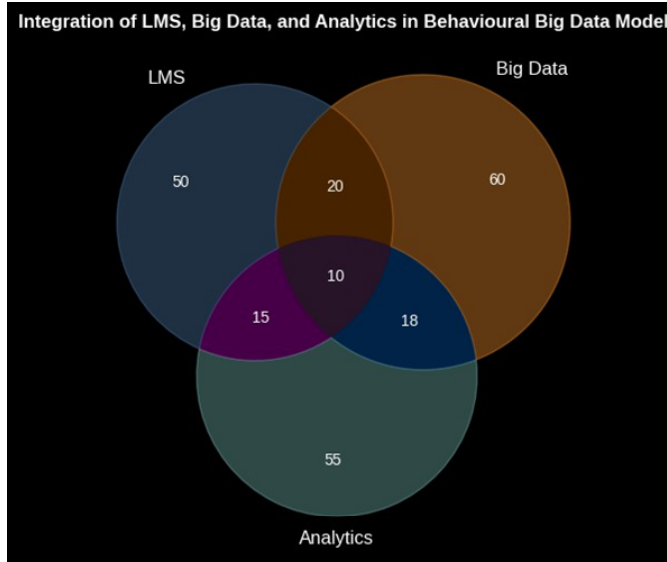

Table 3   Descriptive statistical analysis

Variable	Experimental group sample size ()	Sample size of control group ()	Average value ()	Standard deviation (SD)
Learning time (hours)	50	50	30	5
Academic performance (score)	50	50	70	8
Learning strategy score	50	50	3.5	0.6

The important facts about the test group and the control group are shown in Table 3. Different things were used to decide how to learn, how long to learn, and how well the student did. Table 3 shows all of this. The widths, heights, and numbers of people in each group helped us learn more about how people learn and how well learning works in general. The study group learned a lot in 30 hours. The control group did not say how long it took them to learn. The people who took the test got an average score of 70. We still do not know what the control group's mean number is. It is not yet clear what the control group's average score was for how well they learned. The test group got a 3.5 for how well they learned. This kind of number lets us quickly see the information and get a sense of how each group did. We will use these numbers to start a comparison study and then write up the results. A behavioural big data model is used to evaluate the success of teaching in a data analysis course. The basis of this model is the integration of LMSs, big data databases and analytics. The LMS serves as the major source of data about learning behaviour, including records of student involvement and performance.

Large amounts of heterogeneous data may be managed effectively and efficiently with the help of big data infrastructure, which assures scalability. Furthermore, advanced analytics approaches are used to interpret this data in order to identify patterns of behaviour, forecast the results of learning, and evaluate the efficiency of education. When taken as a whole, these components make it possible to conduct a comprehensive and data-driven review of the course. Figure 9 is a Venn image that shows how LMS, big data, and analytics can be combined in a behavioural big data model. Each circle is an area, and the numbers inside them are different and sometimes combine. There are 50 unique parts in LMS, 60 in big data, and 55 in analytics. The overlaps show parts that are

shared between LMS and big data (20), big data and analytics (18), LMS and analytics (15), and a centre overlap of all three fields with ten elements. Overall, the picture shows how these three types of data come together to support a full behavioural analytics system.

Figure 9 Integration of LMS, big data, and analytics within the behavioural big data model for evaluating the effectiveness of teaching a data analysis course (see online version for colours)

The model achieved strong predictive accuracy across all educational levels, with the highest consistency (91%) observed in secondary school data. Interactive, gamified teaching methods improved engagement among primary students, while project-based learning enhanced performance for secondary and university learners. Behavioural clustering revealed that sustained engagement patterns and teacher feedback loops significantly impacted learning gains. These results confirm that the behavioural big data model can adapt across educational stages, offering a scalable, evidence-based approach to evaluate and improve teaching effectiveness from early schooling to advanced data analysis education.

5   Conclusions

The study shows that using behavioural big data models in data analysis classes could make both the students and the teachers much better at what they do. The LMSs, interviews, and financial records all gather much data. This data can help teachers learn more about how their students behave, what they know, and how their work changes over time. The study says that big data analytics is a better way to predict how kids will grow than older methods. That lets us figure out how to teach better, when to step in, and how to make sure that each student has a unique way to learn. More things do not make sense when you use a lot of data and complex deep learning methods like LSTM together. In

general, this cuts down on the number of false results, which makes learning analytics models more accurate. Having hardware issues made it hard to study as much as I would have liked, especially when I was looking into issues in groups and the area. It was still the case even though things went well. The models that were mentioned should be used with better computer networks so that the results can be checked on larger sets of data. This work should be the starting point for further study that builds on this method. If they were used in more types of schools, big data analytics could also change the way we teach even more. Big data projects are important for coming up with new ideas, helping students do better in school, and leading the digital change of higher education. The study looks at everything and shows how important they are.

Declarations

References

Abdul Bujang, S.D., Selamat, A., Ibrahim, R., Krejcar, O., Herrera-Viedma, E., Fujita, H. and Ghani, N.A.M. (2021) 'Multiclass prediction model for student grade prediction using machine learning', *IEEE Access: Practical Innovations, Open Solutions*, 30 June, Vol. 9, pp.95608–95612.

Al Jallad, K., Aljnidi, M. and Desouki, M.S. (2019) 'Big data analysis and distributed deep learning for next-generation intrusion detection system optimisation', *Journal of Big Data*, Vol. 6, No. 1, p.88.

Asemokha, A., Musona, J., Torkkeli, L. and Saarenketo, S. (2019) 'Business model innovation and entrepreneurial orientation relationships in SMEs: implications for international performance', *J. Int. Entrep.*, September, Vol. 17, No. 3, pp.425–453.

Aslam, R. and Khan, N. (2020) 'Constructive feedback and students' academic achievement: a theoretical framework', *New Horizons*, 1 July, Vol. 14, No. 2, pp.175–198.

Aznar, I., Cáceres, M.P., Trujillo, J.M. and Romero, J.M. (2019) 'Impacto de las apps móviles en la actividad física: Un meta-análisis', *Retos*, Vol. 36, No. 7, pp.52–57.

Breier, M., Kallmuenzer, A., Clauss, T., Gast, J., Kraus, S. and Tiberius, V. (2021) 'The role of business model innovation in the hospitality industry during the COVID-19 crisis', *Int. J. Hosp. Manag.*, 1 January, Vol. 92, p.102723.

Cao, Z. (2022) 'Research on prediction and evaluation method of students' performance in colleges and universities based on deep learning', *Information Record Material*, Vol. 11, No. 5, pp.123–125.

Chowdhury, S.A., Islam, M.A. and Kamal, M.A. (2023) *Unveiling the Potential of Big Data Analytics for Transforming Higher Education in Bangladesh; Needs, Prospects, and Challenges*, 10 October, arXiv preprint arXiv: 2311.10727.

Cui, Y. et al. (2022) 'The influence of big data analytic capabilities building and education on business model innovation', *Frontiers in Psychology*, 18 October, Vol. 13, p.999944.

Cui, Y., Kara, S. and Chan, K.C. (2020) 'Manufacturing big data ecosystem: a systematic literature review', *Robotics Comput. Integr. Manuf.*, Vol. 62, p.101861, DOI: 10.1016/j.rcim.2019. 101861.

Fahd, K. and Miah, S.J. (2023) 'Designing and evaluating a big data analytics approach for predicting students' success factors', *Journal of Big Data*, Vol. 10, No. 1, p.159.

Hadwer, A., Gillis, D. and Rezania, D. (2019) 'Big data analytics for higher education in the cloud era', *2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)*, Suzhou, China, pp.203–207, DOI: 10.1109/ICBDA.2019.8713257.

Han, Y., Wei, S. and Zhang, S. (2015) 'Analysis of online learning behaviour from a tutor perspective: reflections on interactive teaching and learning in the big data era', *Asian Association of Open Universities Journal*, Vol. 10, No. 2, pp.29–48.

Jiménez, D., Sancho, P. and Sánchez, S. (2019) 'Perfil del futuro docente: Nuevos retos en el marco de EEES', *Contextos Educ.*, 25 February, Vol. 23, pp.125–139.

Jose, J., Mana, S.C. and Samhitha, B.K. (2019) 'An efficient system to predict and analyse stock data using Hadoop techniques', *Int. J. Recent Technol. Eng.*, Vol. 8, pp.2277–3878.

Li, M., Zhang, Y. and Huang, Z. (2013) 'MOOC: brewing innovations in higher education', *China Science Daily*, 30 May, p.7.

López-Belmonte, J., Pozo-Sánchez, S., Fuentes-Cabrera, A. and Trujillo-Torres, J-M. (2019) 'Analytical competences of teachers in big data in the era of digitalised learning', *Educ. Sci.*, 10 July, Vol. 9, No. 3, p.177.

Tariq, M.U., Babar, M., Poulin, M., Khattak, A.S., Alshehri, M.D. and Kaleem, S. (2021) 'Human behaviour analysis using intelligent big data analytics', *Front Psychol.*, 6 July, Vol. 12, p.686610, PMID: 34295289, PMCID: PMC8290162, DOI: 10.3389/fpsyg.2021.686610.

Villegas-Ch., W., Mera-Navarrete, A. and García-Ortiz, J. (2023) 'Data analysis model for the evaluation of the factors that influence the teaching of university students', *Computers*, 31 January, Vol. 12, No. 2, p.30.

Wei, S. (2013) 'Learning analytics: mining the value of education data under the big data era', *Modern Educational Technology*, Vol. 23, No. 2, pp.5–11.

Yan, J., Chen, A. and Wang, H. (2024) 'An analysis of English learning behaviour based on big data and its impact on teaching', *Journal of Computational Methods in Sciences and Engineering*, Vol. 24, No. 1, pp.235–251.

Zhang, C. et al. (2024) 'Simulation-based machine learning for predicting academic performance using big data', *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, Vol. 16, No. 1, pp.1–20.