BIG DATA & ARTIFICIAL INTELLIGENCE: STRATEGIC TECHNOLOGY TREND & SMART LEARNING ANALYTICS

By
Salleh Abd Rashid
Zuraidah Mohd Zain
Zul Azhar Zahid Jamal

Background

Thank you for inviting me to deliver this keynote speech. I feel very honoured to be among old and new friends here today.

The topic of my presentation is Big Data & Artificial Intelligence: Strategic Technology Trends & Smart Learning Analytics. It is a huge area to cover, and as I sat down to prepare for the talk, I felt overwhelmed, because it is an area that develops rapidly even as I speak. After much thought, I decided to focus on a narrow aspect of the title – namely, that of Big Data and Artificial Intelligence as used in higher learning institutions. I think most of us here in this gathering belong to the higher learning sector, and I would like to present what is relevant to all if not most of you. I do not intend to make this a highly technical discourse, for, I am not a Big Data expert. I cannot say that I know much about it, because, I am a Physicist by training. However, in the last two decades or so, I held various university administration and leadership posts, which I believe have afforded me a unique perspective to evaluate how Big Data and Artificial Intelligence should be used in the sector so as to acquire maximum benefit to all stakeholders

Let me explain the structure of this presentation. It is divided into 3 parts. The first part is an overview of the concepts that I will be discussing in later parts. This is important so that you are with me, and that we will be on the same page as I deliberate on the concepts later. Part 2 looks into the use of Big Data in higher learning institutions. I will give some examples found in recent literature. In Part 3, I will discuss the use of Big Data and Artificial Intelligence in Malaysian higher learning institutions, particularly in Universiti Malaysia Perlis (UniMAP), where I am from. But I must confess - not much has been published with regards to the use of Big Data in universities, and hence what I will be sharing is

predominantly based on my own experience and what I would like to see happening in UniMAP. The focus here would be UniMAP's attempt at Big Data in learning analytics.
Ladies and gentlemen,

Overview of Terms

As I stated earlier, I will start by going into basic definitions and terminologies that will be used in subsequent sections. Although the terms 'Big Data' and 'Artificial Intelligence' are used a lot these days, it is useful to review them. Sometimes, our understanding of a term changes as we hear or read a lot on it. It's best that we agree on a standard meaning of these essential terms.

So, what is 'Big Data'? As its name implies, Big Data is principally data sets – whether unstructured, semi-structured or structured - that are very large in size. When I say large, I really mean large. Some reports indicate that we create 2.5 quintillion bytes of data every day. Do you know how large that is? It is so large that it requires special technologies and techniques before any sense can be made off the data. Most of the time, Big Data is about unstructured data sets – ones that people tend to overlook, that contain insights that are useful. The examination of data sets may uncover patterns, trends, correlations, and other information that are beneficial to a myriad of organisations in a range of sectors like healthcare, retail, public services, and the hospitality industry. With insights that may otherwise remain unknown, Big Data Analytics help industries make better and more informed decisions. Preventing diseases, battling crimes, and identifying business trends are some benefits that are enabled by Big Data Analytics.

As for 'Artificial Intelligence', it is effectively intelligence exhibited by machines as opposed to intelligence in human beings. Artificial Intelligence is not new. As far back as 1956, people were already talking about the possibility of making machines think like human beings. The era of the construction of the artificial brain had then begun. Learning algorithms were created and fed into computers to enable the process of learning, which then led to the process of thinking. It is through learning algorithms that machines could be made to think.

With the enormous amount of data now available in many unstructured forms and from multiple sources, the human mind cannot cope effectively with the tasks of sorting, classifying, interpreting, and refining the data. This is where Artificial Intelligence comes in. Data is needed for machines to be able to think - the more data, the better. Hence, Artificial Intelligence uses Big Data extracted from the Internet to train algorithms, which then enables thinking to be done by machines. With Artificial Intelligence, data can now be turned into knowledge, which can then be acted upon. To put it in simplistic terms, Big Data and Artificial Intelligence go hand in hand.

Big Data in Higher Learning Institutions

Now that the basic terminologies are out of the way, let me zoom into how higher learning institutions around the world are using Big Data and Artificial Intelligence. My brief survey of the literature finds that the university that takes advantage of Big Data identifies as many qualified candidates as possible, and then work towards making itself **more appealing** to the pool of candidates identified, so that the candidates would choose it over other universities. The data sources used are student test results, social media pages, school cameras, campus sensors, patterns of browsing of particular websites, mobile phone devices, and many more.

Some colleges even provide grants to do research on ways to extract data that can be used to garner actionable insights. The investment of money and time in doing all this is understandable, given the fact that with rising education costs, the number of qualified and financially capable students are getting smaller and smaller, hence forcing universities and colleges to compete with one another to win the hearts of qualified candidates.

Let me describe a bit more on what I mean by ‘qualified candidates’ and what the university does to make itself more appealing to these ‘qualified candidates’.

Say, for example, a college aspires to produce as many academically excellent graduates as possible. One of the traditional ways to do this is to attract academically good students via the disbursement of scholarships, targeting at candidates with the highest academic scores. However, not all scholarship holders would fare well once in university. Hence, there is no guarantee that even after selected for scholarship based on academic performance, the candidate would pass university examinations with flying colours.

Now, with Big Data, the university can increase its chances to confer scholarships to the best possible candidates by analysing more of the candidates' previous academic performances. In other words, the university does not have to depend on the results of one examination only to make a decision on whether or not the candidate is qualified to enrol in the university. Rather, with Big Data, it can now use a collection of information sources from the candidates' many years of schooling, hence allowing for more accurate prediction of the candidates' likely academic performance in university. This way, the likelihood that the university produces more excellent graduates is increased accordingly.

What I have just described is the use of Big Data in the process of student recruitment. Apart from at the entry point, a lot is done with Big Data throughout the student's stay in the institution. More and more universities now use programs and applications to track students' learning habits. Some examples are – how long students spend time online, the types of online websites they frequent, whether they participate in online forums, and how long it takes them to complete their assignments. All this data is studied in order to highlight who requires special attention from lecturers, and what type of special attention is needed in order for the students to stay academically abreast. This will in turn assure a higher probability of academic success.

On the part of the lecturer, Big Data helps by enabling feedback of student performance very often and very quickly. This way, when the lecturer sees a large number of failures, an investigation is performed, which leads to analysis of whether there is a problem with the standard of teaching, or the content of the subject, or whatever else. In other words, lecturers (and university administrators) continuously gauge students' understanding on the academic subject matter, and correspondingly adjust their teaching strategies so as to be more in line and relevant.

Some learning institutions carry this to a higher level. With sufficient data on students' academic behaviour and performance, the university can create a customized learning programme for each and every one, even if there are thousands of students in the university. Thus, every student has the opportunity to study at his/her own pace, therefore making the slower ones acquire knowledge in a more palatable way. In this regard, the university usually goes for 'blended learning', which is a combination of online and offline learning programmes, to fulfil every student's learning needs.

Assessment of curriculum is another area of how the university takes advantage of Big Data. We have all heard about universities offering courses that are not what the industry wants. As a result, there is a mismatch between the graduates that universities produce and the human resource capability that the industry is interested to hire. This is where Big Data comes in handy. The university assesses students' skills and competencies to a level detailed enough such that the industry can narrow down potential employees even when they are still in school. With Big Data in hand, predictive analytics are carried out to give an insight into future student outcomes, which is what potential employers want. With data on predicted future outcomes, employers have ample time working with the university to train selected students so that upon graduation, some selected students can immediately work in the industry without having to be further trained. Put it another way, Big Data allows for selected students to acquire just the right amount of training and education that their potential employers require.

Potential employers are not the only target party to benefit from predictive analytics. In actual fact, the university benefits more, because, if future student outcomes are not satisfactory, the university can adjust its academic programmes appropriately without having to go through the conventional trial-and-error route. Running scenario analysis and acting on the results in terms of curriculum evaluation save the university an enormous amount of money, time, and reputation. In addition, actual data of post-graduation student performance, when fed into the data system, predicts the performance of future graduates, hence helping potential students make informed decisions when choosing the right university.

Apart from student academic files, Big Data also comes in convenient in terms of staff records. Each staff member has a digital footprint that proves advantageous if acted on appropriately. For example, a lecturer's teaching competencies (as seen from his/her students' academic performance, quality of teaching as judged by students, student class attendance, etc.) are linked with other related aspects such as research competencies (as seen from amount and number of research grants acquired, number of papers published, number of conference attended, etc.) and salary, in order to help the university to build the best ecosystem that produces the finest quality of work.

Outside of the academic sphere, Big Data is also used to save on the university's day-to-day operating cost. For example, real-time data obtained from various devices used in the campus like sensors, actuators, and meters, are collated to construct models of energy consumption. These models are then used to track, forecast, and optimise campus energy utilisation. Buildings that have high electricity usage compared to others are uncovered, and the appropriate actions are taken to reduce the usage, and hence reducing electricity bills.

Ladies and gentlemen,

Big Data in Malaysian Higher Learning Institutions and UniMAP

So, what is the scenario like regarding Big Data in Malaysian higher learning institutions? As I said earlier, there is a dearth of published reports on this, and hence, what we know is not as much as what we would like for it to be. But one thing is for sure – information and communication technology (or ICT) has been the government's pillar of strength in ensuring that universities are at the forefront of governmental institutions, and hence billions have been poured into developing the best ICT infrastructure since about two decades ago. With good ICT infrastructure already in place, harnessing what Big Data has to offer should not be too much of a problem.

In UniMAP, we have been benefiting from the use of **staff data** for many years now. For example, we have a range of information of all our staff members – from their personal details such as age, family statistics, and previous academic achievements - to service records such as salary scale, promotion undertakings, number of leaves entitled and taken, research accomplishments, teaching successes and other work attainments. Unfortunately, although all the data are in good and organised order, they are not linked in a way that will enable more informed decision-making to be carried out. Put it another way, we have not quite arrived at a level where insights can be garnered out of this vast amount of data.

Likewise, we also work on a myriad of **student data** – from incoming information comprising accomplishments prior to entry into university, to academic history, economic status, family standing, on-going learning progress, and final academic achievements. Let me elaborate on what I mean by on-going learning progress. In particular, let me focus on e-learning. We started e-learning more than a decade ago, and worked aggressively on it starting about four years ago. Out of more than 1200 courses the university offers, approximately 600 are simultaneously available online – hence making for a blended learning package.

From our e-learning courses, we are able to capture a range of information that proves useful in our deliberation for future improvement. For example, we found that students go online quite often. In one particular week (between 1st May and 7th May 2019), there were 2460 hits. As a matter of fact, almost every day, at least 2000 students access the platform, with hits reaching 4000 on some days. Nevertheless, weekends, particularly Sundays, are slow days.

Interestingly, the data shows that the visits made by the students were only for a short duration of time – mostly between 2 – 3 minutes, with a maximum of 6 minutes only. The time of access was predominantly between 8 am – 10 am in the morning, and 8 am to 10 pm in the evening. The students used mostly desktop (55%) and mobile devices (44%), giving an indication that future development of e-materials should be made to be easily accessed via mobile phones.

What I have shared with you is just a small slice of data that we are actively harnessing and attempting to make sense of. A lot of discussions and decisions have been taken up from these raw data, leading to much improvement in our ecosystem. I can safely say that we have Big Data in place, but what we lack now is the Artificial Intelligence bit of the equation. We have the means to collect more data than what I have described to you, but our interpretation of it is still carried out manually – at least for many of the applications we are using. Hence, although we are very much aware of the benefits that Big Data can offer, we still have a lot of grounds to cover. We are however working on it and are steadily making progress every day.

Fortunately, there are some higher learning institutions in Malaysia that have invested much more than us in this undertaking. Learning analytics, for example, have been used to the extent that student academic achievements are automatically generated every single day, and improvement processes are acted upon immediately. For example, student results from assignments, quizzes and tests are auto-analysed, with action items swiftly displayed for lecturers to take. The exact topics and subject matters that students find difficult are promptly exhibited so that repeat classes or more in-depth tutorials can be offered straight away.

E-learning is a great area where Big Data and Artificial Intelligence can be put to good use. However, it is not without its own challenges. To start with, UniMAP is an engineering university offering engineering courses. Our courses are subjected to accreditation rulings from our local engineering professional body. This professional body, at the moment, does not find e-learning suitable for use in a number of engineering subjects other than just as a supplement to the conventional face-to-face approach. This is understandable, as we are also finding it quite challenging to develop skill-based, as opposed to theory-based subjects, to be used on the e-platform. With this difficulty, there is the issue of credit transfer, which still needs to be sorted out. There is also the issue of access to the internet. Although it is safe to say that almost all students do have access, the fact remains that this number is not 100%. Given the fact that we do not want any student left behind, anything less than 100% is not good enough. So, in other words, while we want to go for e-learning in a huge way, we still need to work on the multitude of challenges facing us.

Closing

Ladies and gentlemen,

I hope I have succeeded in giving you a picture of how Big Data and Artificial Intelligence can be of great use to higher learning institutions, and how, in reality, it is actually used in UniMAP. There is a lot of potential, but much expenditure needs to be made too. The universities that are ahead of UniMAP have invested millions of ringgit to get to where they are. In the economic climate we are facing right now, we cannot afford to put in as much money as we would like to. So until then, we have to be contend with making the best of what we already have, but at the same time, we prepare for opportunities that will enable us to bring Big Data and Artificial Intelligence to a higher level.

Thank you.