

## Industrial Internship Report on

## ” Online Education Platform ”

Prepared by

Smitha S V

### *Executive Summary*

This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT).

This internship was focused on a project/problem statement provided by UCT. I had to finish the project including the report in 6 weeks' time.

My project was that, we are working with the government to transform various cities into smart cities. The vision is to convert it into a digital and intelligent city to improve the efficiency of services for the citizens. One of the problems faced by the government is education. You are a data scientist working to manage the online education better and to provide bright future.

The government wants to implement a robust education system for the city by being prepared for online education platform. They want to understand the education patterns that is currently happening in the city.

## **TABLE OF CONTENTS**

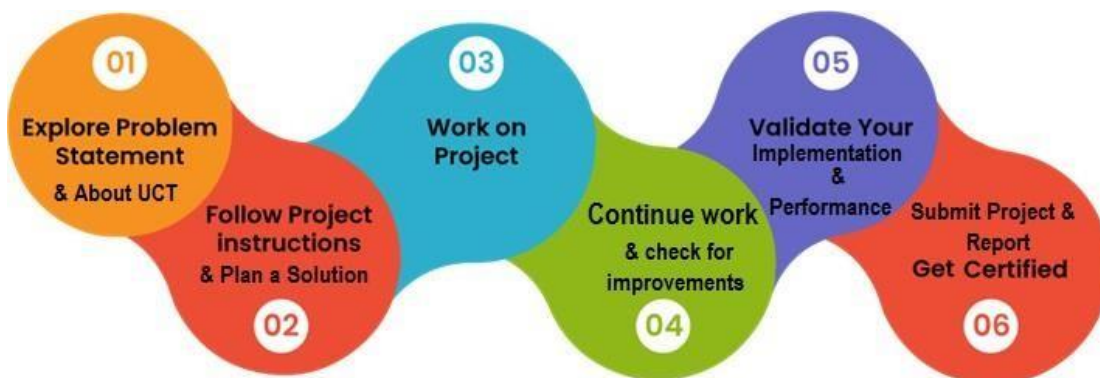
1	Preface.....	3
2	Introduction.....	5
2.1	About UniConverge Technologies Pvt Ltd .....	5
2.2	About upskill Campus .....	9
2.3	Objective.....	10
2.4	Reference .....	10
2.5	Glossary .....	
3	Problem Statement.....	11
4	Potential solution .....	12
5	Development of New Model.....	15
5.1	Research Model and Hypothesis.....	
5.2	Measurement Model.....	
6	Evaluation of online performance.....	22
6.1	General form of decision .....	25
7	My learnings.....	31
8	Future Work scope.....	32

## 1 Preface

In terms of population and economic growth, India is the second-largest country. India's major cities are plagued by issues with education system. In wealthy nations, sustaining Education Systems is a realistic challenge, and enhancing urban areas is also fraught with challenges. One of the most significant issues facing India is this one. This is brought on by limited space, high costs, and delayed infrastructure growth compared to the quick rise in population. People need free flow of information to aid in better education in all possible aspects and to increase the effectiveness in obtainment of knowledge. To manage knowledge networks, it will be more beneficial to forecast short-term courses in online platforms. These short-term courses are useful to increase the knowledge and improve the technologies in better ways possible. Every nation on earth strives to make their education system more effective. Different techniques have been employed by researchers for betterment of education system. The most prevalent issue in many of the emerging cities, despite the growth of transit services, is education. With the fast expansion of the IT industry, both the population and the number of people to educate in major cities like Bangalore and Delhi. To provide education, online platform is the best way. It saves the travel time; a smart and intelligent education system is needed. This is why India's urban regions need online education platform.

Cloud computing means storing and accessing the data and programs on remote servers that are hosted on the internet instead of the computer's hard drive or local server. Cloud computing is also referred to as Internet-based computing.

**Cloud computing** is adopted by every company, whether it is a MNC or a startup and many are still migrating towards it because of the cost-cutting, lesser maintenance, and the increased capacity of the data with the help of servers maintained by the cloud providers. One more reason for this drastic change from the On-premises servers of the companies to the Cloud providers is the **‘Pay as you go’** service provided by them i.e., you only have to pay for the service which you are using. The disadvantage On-premises server holds are that if the server is not in use the company still has to pay for it.



## 2 Introduction

### 2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies** e.g., **Internet of Things (IoT)**, **Cyber Security**, **Cloud computing (AWS, Azure)**, **Communication Technologies (4G/5G/LoRaWAN)**, **Java Full Stack**, **Front end** etc.



#### i. UCT IoT Platform ( **Insight** )

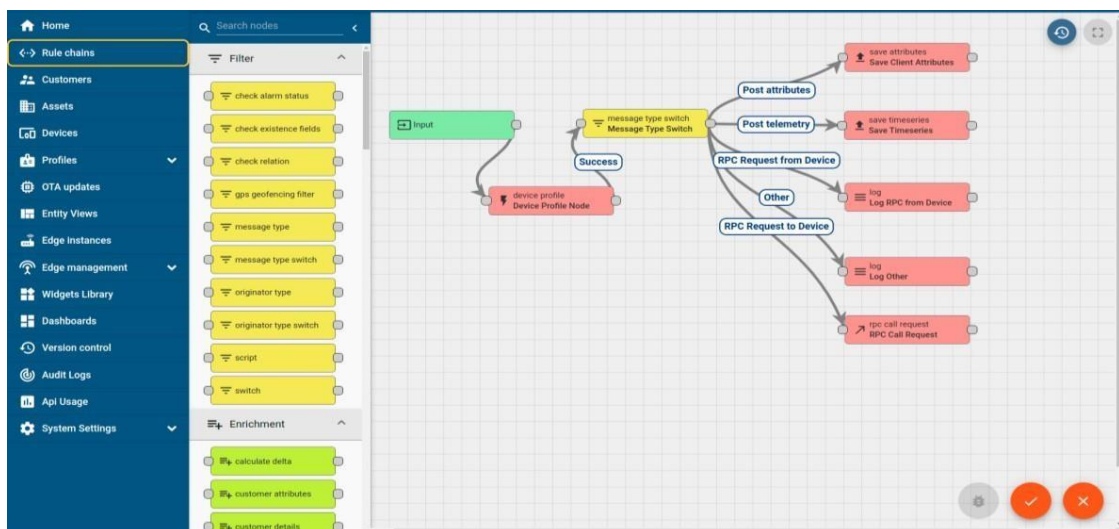
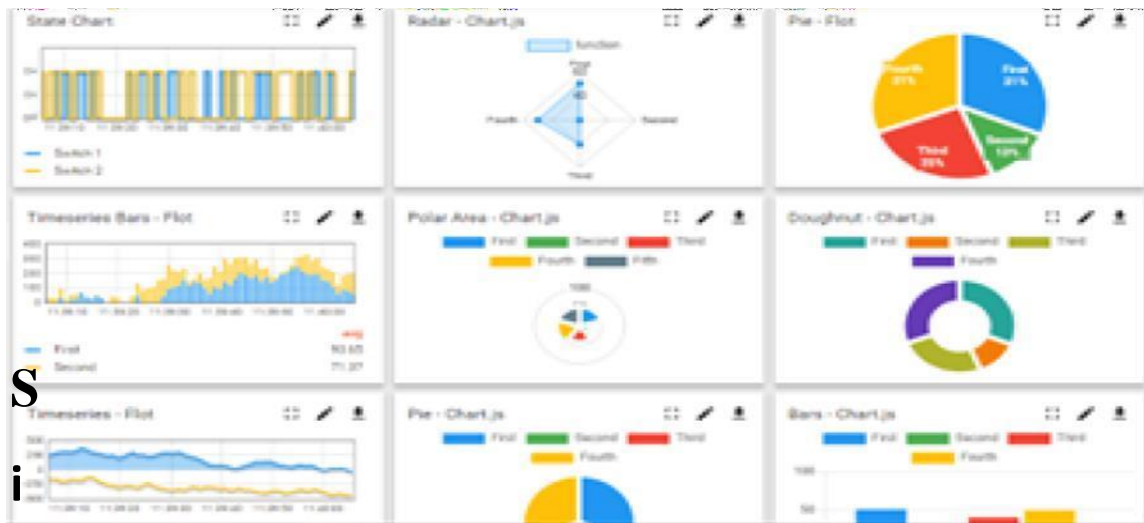
**UCT Insight** is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSQL Databases.



- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

It has features to

- Build Your own dashboard
- Analytics and Reporting
- Alert and Notification
- Integration with third party application (Power BI, SAP, ERP)
- Rule Engine



## FACTORY WATCH

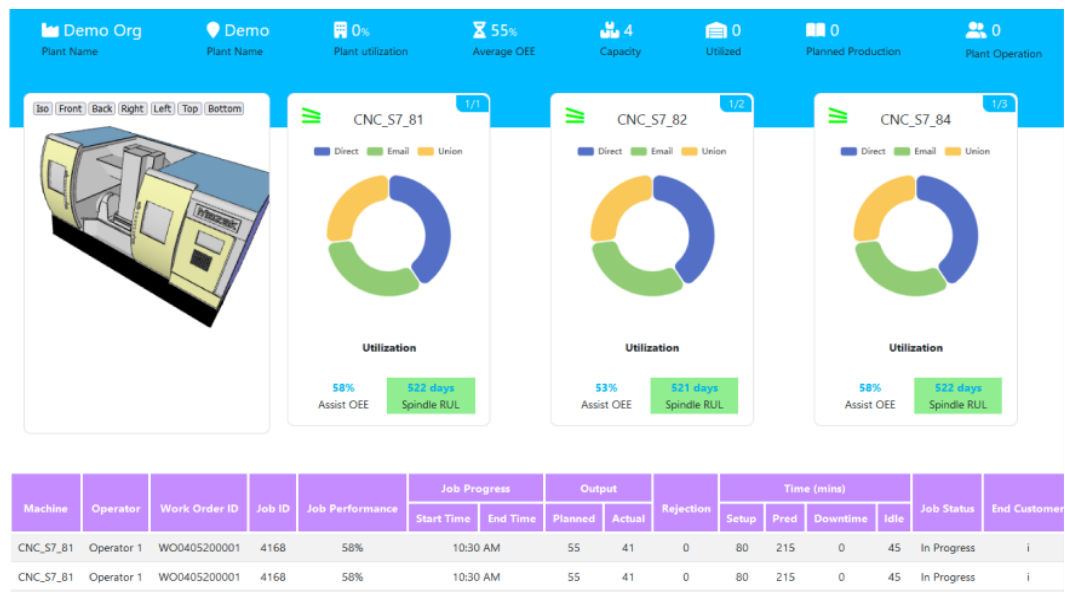
### ii. Smart Factory Platform ( )

Factory watch is a platform for smart factory needs.

It provides Users/ Factory

- with a scalable solution for their Production and asset monitoring
- OEE and predictive maintenance solution scaling up to digital twin for your assets.
- to unleashed the true potential of the data that their machines are generating and helps to identify the KPIs and also improve them.
- A modular architecture that allows users to choose the service that they what to start and then can scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.





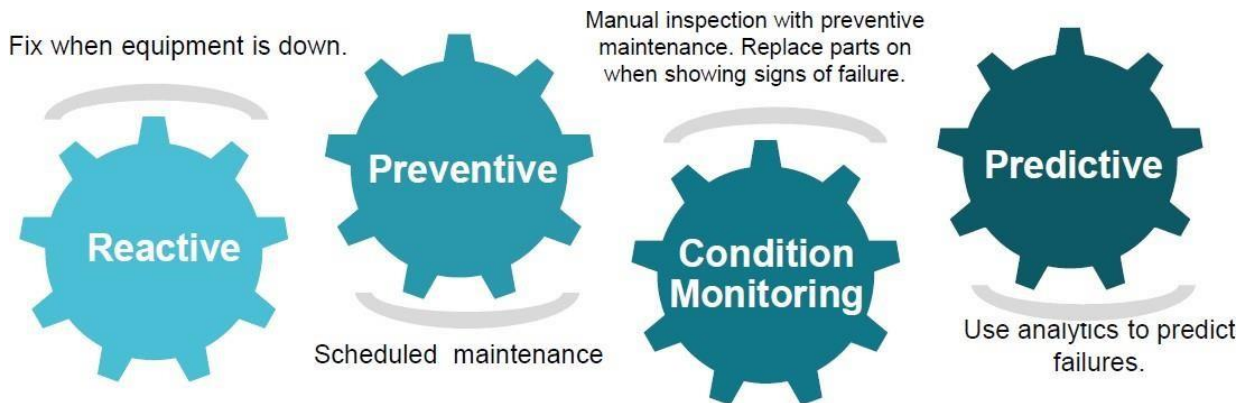
### iii. based Solution

UCT is one of the early adopters of LoRAWANteschnology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

### iv. Predictive Maintenance

UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.

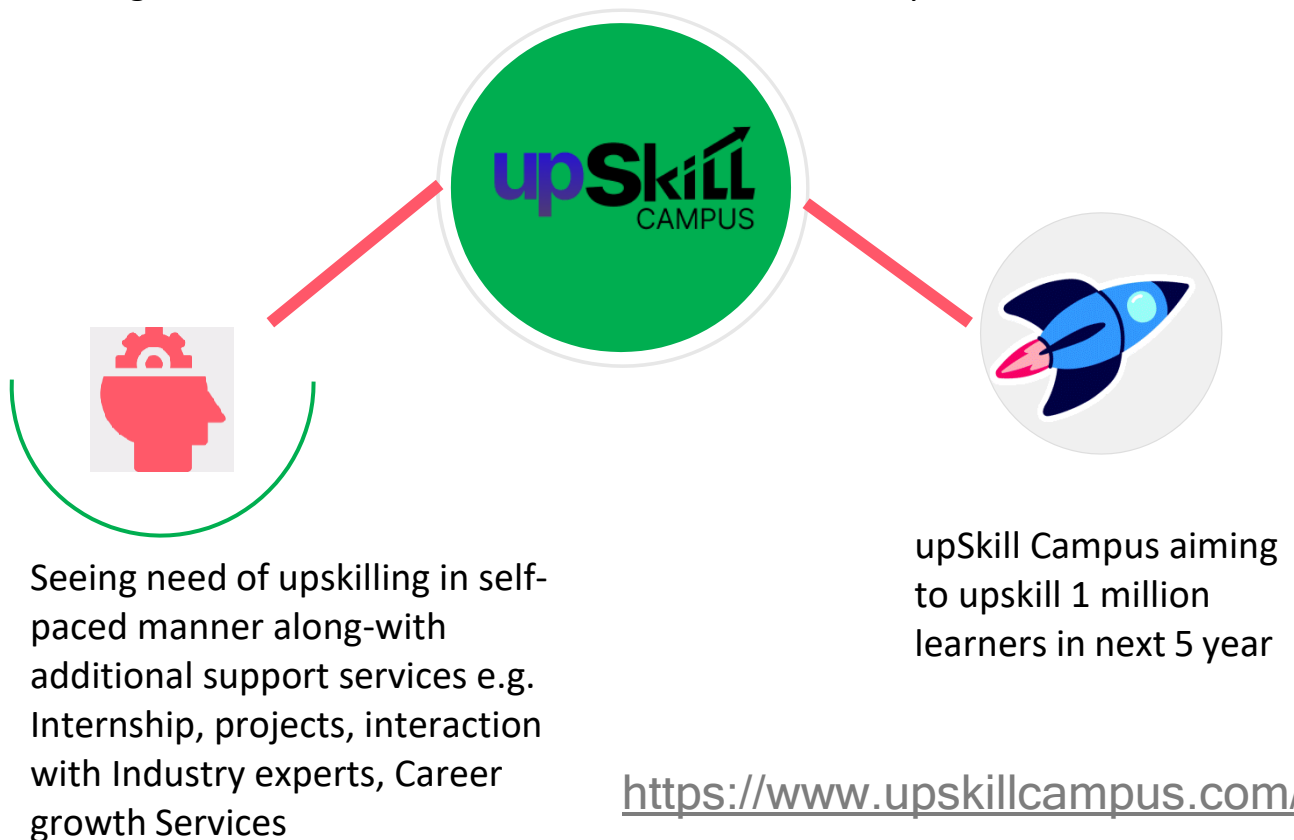


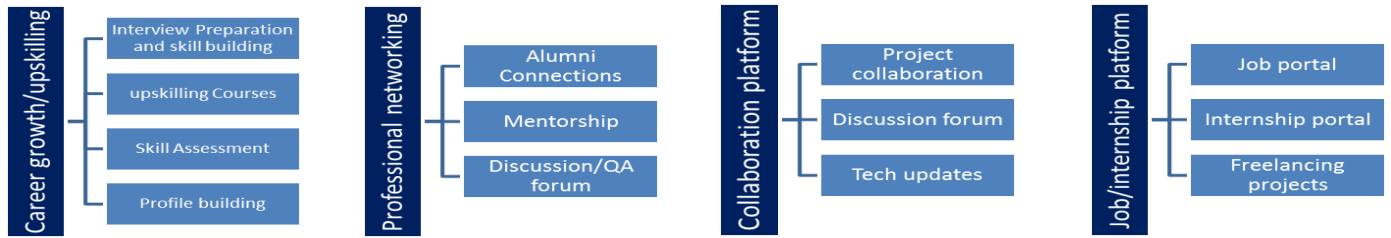


## 2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.





## 2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs in collaboration with EICT Academy, IITK, IITR and IITG in multiple domains.

## 2.4 Objectives of this Internship program

The objective for this internship program was to

- get practical experience of working in the industry.
- to solve real world problems.
- to have improved job prospects.
- to have Improved understanding of our field and its applications.
- to have Personal growth like better communication and problem solving.

## 2.5 Reference

- [1] <https://learn.upskillcampus.com/s/mycourses>
- [2] <https://www.uniconvergetech.in/>
- [3] <https://upskillcourses.com/>

### 3 Problem Statement

Online learning platforms have revolutionized the way education is delivered, especially in the wake of the COVID-19 pandemic. However, they also present a unique set of challenges that can hinder effective learning. Here is a problem statement that encapsulates these issues:

"Despite the convenience and accessibility of online learning platforms, their use for educational purposes is often hampered by issues such as lack of personal interaction, technical difficulties, limited access to resources, and the need for self-discipline and motivation. These challenges can lead to decreased student engagement and learning outcomes, creating a need for improved strategies and tools to enhance the online learning experience."

#### Key Problems

1. **Lack of Personal Interaction:** Online platforms often lack the face-to-face interaction found in traditional classrooms, which can lead to feelings of isolation and decreased student engagement.
2. **Technical Difficulties:** Students and teachers may face technical issues such as unstable internet connections, software glitches, and lack of familiarity with the platform.
3. **Limited Access to Resources:** Not all students have equal access to the necessary technology or internet connectivity, leading to disparities in learning opportunities.
4. **Need for Self-Discipline and Motivation:** Online learning requires a high degree of self-discipline and motivation, which can be challenging for some students.
5. **Decreased Learning Outcomes:** The aforementioned issues can lead to decreased learning outcomes, as students may struggle to engage with the material and retain information.

#### 4 Potential Solution

Addressing these problems requires a multi-faceted approach:

- Enhancing the **interactive capabilities** of online platforms, such as video conferencing and discussion forums.
- Providing **technical support** and training for students and teachers.
- Ensuring **equal access** to necessary resources through initiatives like loaner programs or subsidized internet access.
- Implementing strategies to **boost student motivation** and self-discipline, such as regular check-ins and feedback.
- Continually **assessing and improving** the effectiveness of online learning through surveys and data analysis.

By addressing these issues, we can work towards making online learning a more effective and equitable tool for education.

Like in many countries, in the fight against COVID-19, schools in North Macedonia have closed and learning is taking place online. The amount of content, online assessment, much of it without adequate instructions, are just some of the problems that young people face in this new learning environment. The Ministry of Education and Science in cooperation with UNICEF and other partners created a new online-learning platform called EDUINO, where pre-primary and primary school students can continue their education through video lessons, resources and a variety of games. In addition an educational program called TV Classroom is being broadcast on the national television. The primary and secondary schools have also started holding online classes through various digital platforms. All of these are good efforts, but - ***do all students have the means to attend online classes and what can we say about the actual quality of education in this online environment?*** These are some different perspectives and views on how we can make it work.



## Student's Perspectives

The truth is, for many students, online learning is only a formality and not a real substitute for regular teaching. Some teachers only share material to students without teaching it. Online testing is sometimes based on the principle of “work it out yourself”. Students are not acquiring real, long-lasting knowledge. And some students don't have the opportunity to leave their home during the two hours allowed during the curfew because they have to sit in online classes. Some students don't even have proper equipment to attend online classes. They don't have electronic devices such as computers, telephones and cameras. The number of these devices in households is often limited which can be very inconvenient for online appointments, classes, and meetings that take place simultaneously. Also, some teachers don't consider the fact that during online testing, the student may lose the internet connection. Unfortunately, if this happens, the student gets graded based on the number of questions answered and recorded in the system before the connection was lost. Students also face problems managing their own time as a result of online teaching.

Reference :

[\[Solved\] what is a problem statement for using online learning platforms - Inclusive Education \(#12345\) - Studocu](#)

### 4.1 Code submission ( GitHub link)

<https://github.com/SmithaSV/Onlineeducationplatform.git>

### 4.2 Report submission (GitHub link): first make placeholder, copy the link.

## Teachers Perspectives

For now, everyone is going on as if the most important thing is to teach what is the remaining curriculum, to get the final grades and to finish the school year formally. But is it really necessary for students? Is that the right way to deal with this new situation?

Certainly not! In this big picture, perhaps the biggest burden is put on teachers. They are in a situation where they are unprepared and without proper support. Criteria and guidelines imposed by the institutions are not sufficient to deal with the situation effectively. Existing assessment criteria that include tests and examinations are not suitable for digital learning. No teacher can assess with certainty whether the homework assigned to students is written independently and assigning separate homework to each individual student is simply an overload and difficulty.

Teachers need serious preparation to use online tools and platforms. They are not all ready for the new situation, which further opens the issues with our overall education. We are all aware that if we want to improve the quality of education, we need to better use digital technologies, but we also need to provide appropriate support and training to teachers to support the quality of instruction.

## 5 Development of new Model

This research aims to explore and investigate potential factors influencing students' academic achievements and satisfaction with using online learning platforms. This study was constructed based on Transactional Distance Theory (TDT) and Bloom's Taxonomy Theory (BTT). This study was conducted on 243 students using online learning platforms in higher education. This research utilized a quantitative research method. The model of this research illustrates eleven factors on using online learning platforms to improve students' academic achievements and satisfaction. The findings showed that the students' background, experience, collaborations, interactions, and autonomy positively affected students' satisfaction. Moreover, effects of the students' application, remembering, understanding, analyzing, and satisfaction was positively aligned with students' academic achievements. Consequently, the empirical findings present a strong support to the integrative association between TDT and BTT theories in relation to using online learning platforms to improve students' academic achievements and satisfaction, which could help decision makers in universities and higher education and colleges to plan, evaluate, and implement online learning platforms in their institutions.

### INTRODUCTION

Higher education organizations over the previous two decades have offered full courses online as an integral part of their curricula, besides encouraging the completion throughout the online courses. Additionally, the number of students who are not participating in any courses online has continued to drop over the past few years. Similarly, it is perfectly possible to state that learning online is obviously an educational platform. Courses online are trying to connect social networking components, experts' content, because online resources are growing on daily basis.

Such courses depend on active participation of a significant number of learners who participate independently in accordance with their education objectives, skills, and previous background and experience. Nevertheless, learners differ in their previous background and experience, along with their education techniques, which clearly influence their online courses results besides their achievement. Consequently, despite the online learning evolution, learning online possibly will not be appropriate for each learner.

Learners' satisfaction and academic achievement towards learning online attracted considerable attention from scholars who employed several theoretical models in order to evaluate learners' satisfaction and academic achievements. This present study highlights the effects of online learning platforms on student's satisfaction, in relation to their background and prior experiences towards online learning platforms to identify learners that are going to be satisfied toward online course. Furthermore, this research explores the effects of transactional distance theory (TDT); student collaboration, student- instructor dialogue or communication, and student autonomy in relation to their satisfaction. Accordingly, this study investigates students' academic achievements within online platforms, utilizing Bloom theory to measure students' achievements through four main components, namely, understanding, remembering, applying, and analyzing. This study could have a significant influence on online course design and development. Additionally, this research may influence not only academic online courses but then other educational organizations according to the fact that several organizations offer training courses and solutions online. Both researchers and Instructors will be able to utilize and elaborate in accordance with the preliminary model, which was developed throughout this research, on the effects of online platforms on student's satisfaction and academic achievements. Advantages of online learning and along



with its applications were mentioned in earlier correlated literature .However, despite the growing usage of online platforms, there is a shortage of employing this technology, which creates an issue in itself. Consequently, the research problem lies in the point that a model needs to be created to locate the significant evidence based on the data of student's background, experiences and interactions within online learning environments which influence their academic performance and satisfaction. Thus, this developed model must be as a guidance for instructors and decision makers in the online education industry in terms of using online platforms to improve students learning experience through online platforms.

### **Research theory hypothesis development**

When designing web-courses within online learning instructions or mechanisms in general, educators are left with several decisions and considerations to face, which accordingly affect how students experience instruction, how they construct and process knowledge, how students could be satisfied through this experiment, and how web-based learning courses could enhance their academic achievements. In this study, we construct our theoretical framework according to Moore transactional distance theory (TDT) to measure student's satisfaction, in addition to Bloom theory components to measure students' academic achievements. Though the origins of TDT can be traced to the work of Dewey, it is Michael Moore who is identified as the innovator of this theory that first appeared in 1972. In his study and development of the theory, he acknowledged three main components of TDT that work as the base for much of the research on DL. Also, Bloom's Taxonomy was established in 1956 under the direction of educational psychologist to measure students' academic achievement (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956). TDT theory has been selected in this study since Transactional distance's term indicates the geographical space between the student and instructor. Based on the

learning understanding, which happens through learner's interaction with his environment. This theory considers the role of each of these elements (Student's autonomy, Dialogue, and class structure) whereas these three elements could help to investigate student's satisfaction. Moore's (1990) notion of 'Transactional Distance' adopt the distance that happens in all relations in education. The distance in the theory is mainly specified the dialogue's amount which happens between the student and the teacher, and the structure's amount in the course design. Which serves the main goal of this study as to enhance students online learning experience in relation to their satisfaction. Whereas, Bloom Theory has been selected in this study in addition to TDT to enhance students online learning experience in relation to their student's achievements. In a conclusion both methods were implemented to develop and hypothesis this study hypothesis.

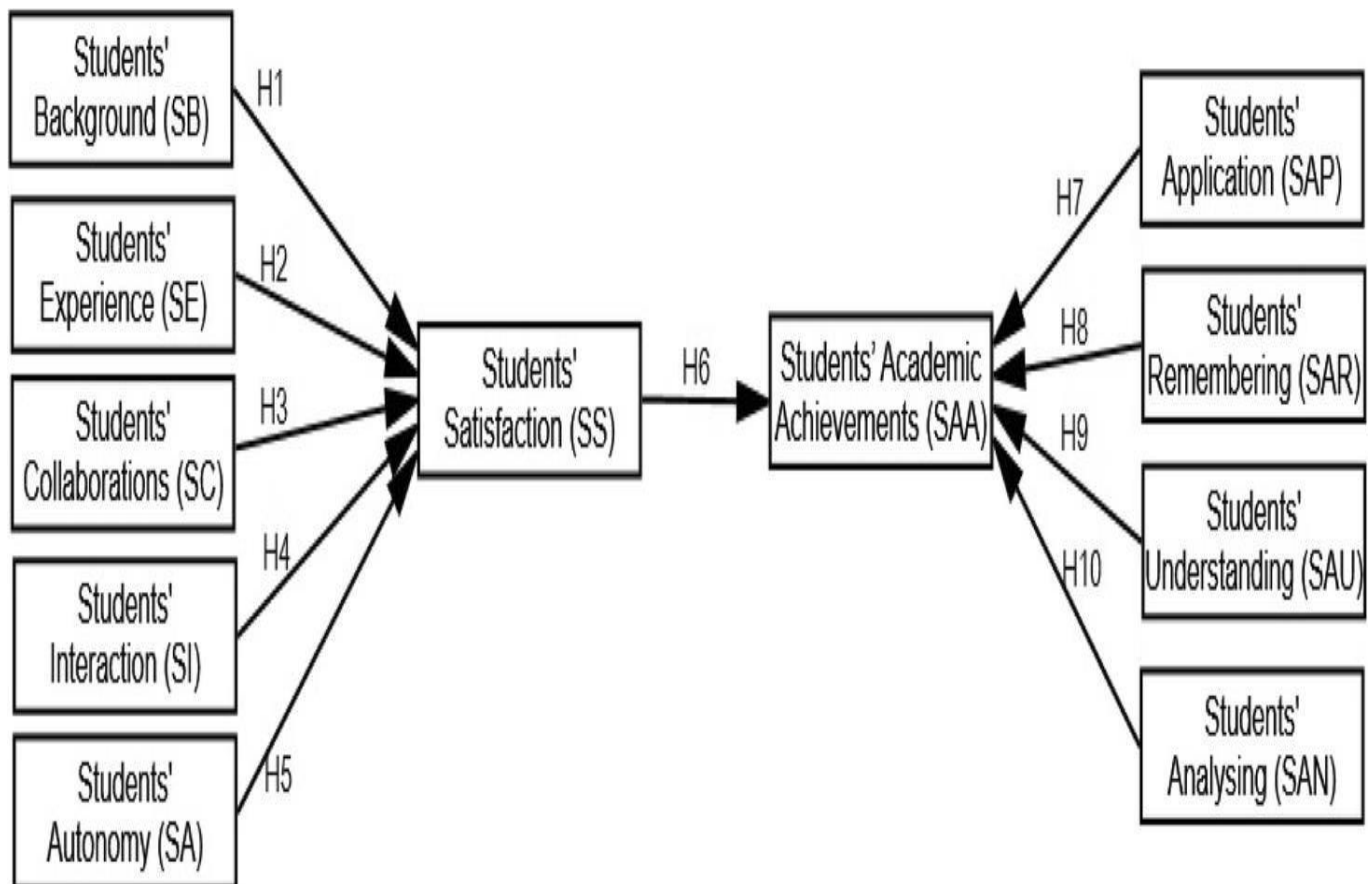


Fig 5.1 : Research Model and Hypothesis

## Students background toward online platforms

Students' background regarding online platforms in this study is referred to as their readiness and willingness to use and adapt to different online platforms, providing them with the needed support and assistance. Students' background towards online learning is a crucial component throughout this process, as prior research revealed that there are implementation issues, for instance; the deficiency of qualified lecturers, infrastructure and facilities, in addition to students' readiness, besides students' resistance to accept online learning platforms in addition to the Learning Management System (LMS) platforms, as educational tools. However, student demand continued to increase, spreading to global audiences due to its exceptional functionality, flexibility and eventual accessibility. There have been persistent apprehensions regarding online learning quality compared with traditional learning settings. In their research, have discovered that Austrian learners continue to prefer traditional learning environments due to communication goals, along with the interpersonal relations preservation.

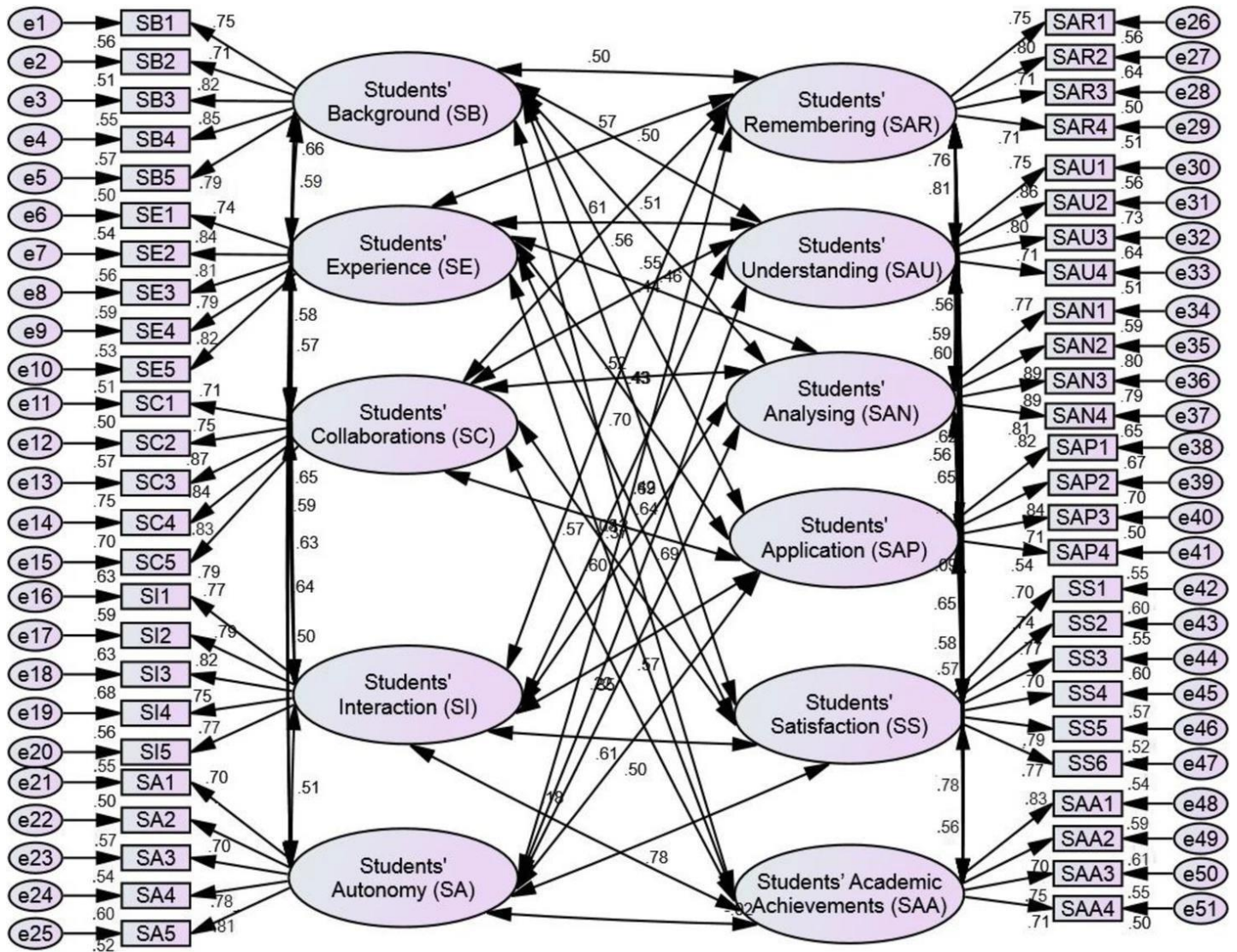
Moreover, have discovered that Malaysian learners' internet efficiency and computer skills, along with their personal demographics like gender, background, level of the study, as well as their financial income lead to a significant difference in their readiness towards online learning platforms. It is claimed that the current technologies in education play an essential role in providing a full online learning experience which is close enough to a face-to-face class in spite of the physical separation of the students from their educator, along with other students. Platforms of online learning lend themselves towards a less hierarchical methodology in education, fulfilling the learning desires of individuals which do not approach new information in a linear or a systematic manner. Platforms of online learning additionally are the most suitable ways for autonomous students.

## Students experience toward online platforms

Students' experience in the current research indicates that learners must have prior experience in relation to utilizing online learning platform in their education settings. Thus, students experience towards online learning offers several advantages among themselves and their instructors in strengthening students' learning experiences especially for isolated learners. Regardless of student recognition of the advantages towards supporting their learning throughout utilizing the technology, difficulties may occur through the boundaries about their technical capabilities and prior experiences towards utilizing the software itself from the perspective of its functionality. As demonstrated over learner's experience and feedback from several online sessions over the years, this may frequently become a frustration source between both learners and their instructors, as this may make typically uncomplicated duties, for instance, watching a video, uploading a document, and other simple tasks to be progressively complicated for them, having no such prior experience. Furthermore, when filling out evaluations, for instance, online group presentations, the relatively limited capability to communicate face-to-face then to rely on a non-verbal signal along with audience's body language might be a discouraging component. Nonetheless, the significance of being in a position to participate with other colleagues employing online sessions, which are occasionally nonvisual, for instance; teleconference format is a progressively significant skill in the modern workplace, thus affirming the importance of concise, clear, intensive interactions skills.



**Fig 5.2 : Measurement model**



## 6 Evaluation Of Online Education Performance

With the continuous development of “Internet + Education”, online learning has become a hot topic of concern. Decision is an important technique for solving classification problems from a set of random and unordered sets. Decision is not only an effective method to generate classifier from set, but also an active research field in technology. The decision can classify the data, grasp the teaching process of the teacher, and analyze the overall performance of the students, so as to realize the dynamic management of the educational administration and help the educational administration personnel to make the right decision, with more reasonable allocation of resources. This paper evaluates students’ academic performance based on the learning behavior data of online learning, so as to intervene in students’ learning in advance, which is the key problem that needs to be solved at present. Taking students’ learning attitude, completion of homework, and attendance as factors, the paper uses decision tree technology to analyze the factors affecting students’ performance, and evaluates students’ performance. Firstly, this paper collects the high-dimensional behavioral characteristic data of students’ online learning and conducts correlation analysis after preprocessing the behavioral characteristics. Then, the decision C4.5 algorithm is used to construct a performance evaluation model. Students’ performance is evaluated by the model, and the evaluation accuracy is about 88% compared with actual performance.

Finally, through the model analysis, it is concluded that the video task point completion is the most influential in students’ achievement, followed by chapter test completion and chapter test average score, and the course interaction amount and homework average score are the least influential in students’ achievement, which has a practical reference value for effectively serving online learning and teachers’ teaching.

## INTRODUCTION

By January 2021, the number of online education users in China reached 452 million, accounting for 49.8% of the total number of Internet users, an increase of nearly 200 million compared with January 2020. The 2020 Government Work Report points out that the government will strengthen and improve the construction of the “Internet + Education” model and promote the equitable development and quality improvement of education, which reflects the government’s determination to strongly support online education. Obviously, online education has become an important part of the development of education. With the expansion of the users of online education platform, a large amount of learning behaviour data has been generated in the background. How to apply these to performance evaluation has become a hot topic of current research. The key purpose of this study is to adopt reasonable algorithm, establish scientific performance evaluation model, and mine learning behaviour data. Based on the analysis results of the model, the key factors affecting students’ online learning are obtained, which provide a theoretical basis for the feedback and guidance of students’ online learning activities, the intervention of teachers, and the implementation of educational evaluation by educational administrators. It is of great significance to guide students to learn online more effectively.

Achievement is an important standard to measure students’ performance and an important basis in the process of teaching quality evaluation. The process of teaching is also a process of data accumulation. By using relevant technologies to analyse, we can fully analyse the rules contained in the performance, carry out quantitative and qualitative analysis of the results, and accurately analyse various aspects of the results according to different students’ learning

conditions. The use visualization technology can not only show the results intuitively but also more clearly show the influence and relationship between various disciplines. According to the relationship between these factors, the educational administration department of the school can modify the teaching plan pertinently, improve the teaching quality of the school, and ensure the results and effects of teaching.

The biggest advantage of cloud computing is to be able to make a comprehensive analysis of a very large amount and to extract rules that can arouse the interest of relevant personnel and some unknown and potential knowledge that is very beneficial to decision-making. Using the tree structure to express the results, users can better understand the relationship between various factors. With the development of visualization technology, artificial intelligence technology, and machine learning technology, many cloud computing methods have been widely used in real life. Achievement, as important data of educational administration in higher vocational colleges, can reflect a student's learning situation and can also evaluate the teachers' teaching quality. The decision method enables students to show the connection between different professional courses and practices through various professional achievements and also helps teachers improve teaching methods. Therefore, in teaching work, more targeted teaching methods can be used to improve the learning efficiency of students and the teaching level of teachers.



## 6.1 General Form of Decision

The decision learning is usually a process of recursively selecting the optimal feature and segmenting the training according to the feature, so as to have the best classification of each sub set. This process corresponds to the division of feature space and the construction of decision. To begin, build the root node and place all the training at the root node. Select an optimal feature, and divide the training set into subsets according to this feature, so that each subset has a best classification under the current conditions. If these subsets can be classified basically correctly, then build leaf nodes and divide these subsets into corresponding leaf nodes. If there are subsets that cannot be classified basically correctly, then select new optimal features for these subsets, continue to segment them, and build corresponding nodes. This process continues until all subsets of training are classified roughly correctly, or there are no appropriate features. Finally, each subset is assigned to a leaf node, so you have a specific class. This generates a decision.

It can be seen from the above process that the generation of decision is a recursive process. In the basic of the decision, there are three scenarios that result in recursive returns: the samples of the current node all belong to the same category and need not be divided. The current attribute set is empty, or all samples have the same value on all attributes and cannot be partitioned. The sample set contained in the current node is empty and cannot be partitioned. In the second case, we mark the current node as a leaf node and set its category as the category where the node contains the most samples. In the third case, the current node is also marked as a leaf node, but its category is set to the category whose parent node contains the most samples. In the second case, the posterior distribution of the current node is used, while in the third case, the sample distribution of the parent node is taken as the prior distribution of the current node.

The decision generated by the above method may have a good classification ability for training but may not have a good classification ability for unknown test ; that is, overfitting phenomenon may occur. We need to prune the generated from the bottom up to make the simpler and thus more generalizable. Specifically, it removes the over segmented leaves, regresses them back to the parent or even higher nodes, and then changes the parent or higher nodes to new leaves. If the number of features is large, the features can also be selected at the beginning of decision tree learning, leaving only those features with sufficient classification ability for training.

It can be seen that the decision learning includes feature selection, decision generation, and decision pruning process. Since the decision represents a conditional probability distribution, the different depth of the decision corresponds to the probability model of different complexity. The generation of the decision corresponds to the local selection of the model, and the pruning of the decision corresponds to the global selection of the model. The generation of decision considers only local optimum, and the pruning of decision considers global optimum.

Decision model has a structure. In classification problem, it represents the process of classifying instances based on features. It can be considered as a set of if-then rules, or as a conditional probability distribution defined in feature space and class space. The classification has the advantages of good readability and fast classification speed. When training the classification, the training data are used to establish the classification model according to the principle of minimizing the loss function. When forecasting, the new data are classified by the

classification model. Decision learning usually includes three steps: feature selection, decision generation, and decision pruning.

A decision can be viewed as a collection of if-then rules: a rule is constructed from each path from the root of the decision to the leaf, where the characteristics of the internal nodes correspond to the conditions of the rule, and the classes of the leaf nodes correspond to the conclusion of the rule. The path to a decision or its corresponding set of if-then rules has an important property: they are mutually exclusive and complete. That is, each instance is covered by one path or one rule. Here, the so-called coverage refers to the conditions that the features of the instance are consistent with those on the path or that the instance satisfies the rules.

- (1) The content of decision has a lot of different methods after a long time of development, usually used in ID3, C4.5, CART, etc. This method will form a decision after cloud computing, through the use of the decision, to help users to make the right decision. In the process of tracking the generation of the decision, it needs to go through the process of decision pruning and decision generation. This can process the unknown and evaluate the future development direction of an item.
- (2) Principle of decision will use certain rules to analyse and sort out during cloud computing, which is widely used in actual life and production. After this pre-processed, the induction method can be used to generate the decision, the related classification work can be carried out, and then the further analysis can be conducted. This method can be used in project risk assessment to analyze the feasibility of the project. This also uses machine learning techniques to analyze the relationships between different values to evaluate and extrapolate the future. A decision is composed of nodes and forks. The nodes represent various

objects, the forks represent various attributes, and the paths passed by each leaf stage represent the attribute values of objects. Forks are generally expressed in the plural form, allowing for relatively independent processing. Therefore, decision can be used for both evaluation and analysis. Decision can conduct in-depth analysis of all the features of the sample, so as to find out the features with decisive significance, then display the analysis results, determine the most significant feature as the root node of the entire decision, and then proceed to analyze the significance of other features to build an inverted . The decision can also be specially designed for nonnumerical, especially for students' grades. Since there are not only numerical but also a lot of nonnumerical, the nonnumerical can be processed according to the characteristics of students' grades. The processing model of the decision is shown in Figure

(3) Advantages of decision :

- (a) It is conducive to users' understanding of rules. Normally, academic administrators conduct analysis on the results of students. As the educational administrators do not have a high level of understanding of cloud computing technology, they need to ensure the interpretation results of cloud computing. This can classify the structure of the decision and generate "If form", which can avoid the difficulty of educators to understand.
- (b) The amount of calculation is relatively small. The analysis of the educational administration system must have strong practicability, which must ensure that the analysis has higher efficiency. In cloud computing, compared to the calculation is relatively small, making the speed of cloud computing much higher than that of others, which can greatly reduce the use of cloud computing time, resulting in very high work efficiency.
- (c) Discrete and continuous process can be processed. There are many different types of students' academic performance, including grades and teachers'

evaluation of students' grades. Therefore, there are not only continuous but also discrete in grades. At the same time, discrete constitutes the majority. For example, many courses adopt discrete variables of high, medium, and low levels, and decision type can process both kinds.

(d) It shows the importance of attributes. The expression way of decision is very intuitive, which can directly reflect the importance of the zodiac, which is directly represented by the current level of decision node. For example, a more important attribute will be at a higher level, and a less important attribute will be at a lower level.

(4) The structure of the cloud computing steps: The cloud computing to construct the decision is divided into two steps, respectively, decision generation and pruning; The decision is a root node that is initially generated and has undergone a gradual process generation from top to bottom. The process of segmentation and formation completes the structure of the decision. Using computing in the decision process, the decision can be tested from the root node, and the decision result will determine the next one until the last point. The pruning of decision is to prune the decision and remove the redundant branches through testing, so as to get the decision with the least expected error rate.

(5) ID3: The currently most commonly used cloud computing is the ID3. The basic idea of the method is recursive downward search on the training sample set. It is a kind of typical greedy. The decision of every node can test each attribute and the use of information gain as the attribute selection criteria and select the maximum attribute of information gain as the node of decision. To build a decision, the concept of information will be used in the to complete.

(6) Information entropy. Information entropy is the expectation of all kinds of information, which can measure the uncertainty of the whole information



element  $X$ . In the set with  $X$  as the sample, all possible signals in the signal source are represented by the symbolic number  $N$ . Let the possible value be  $A_i$ , and the probability when the value is  $A_i$  is  $P(A_i)$ .

- (a) The relationship between information entropy is obvious.
- (b) Conditional entropy: it is the expected value of information entropy under different conditions, where the signal source is by corresponding to  $Y$ , and the signal source is  $a_i$  corresponding to  $X$ , the probability of which is  $p$ .
- (c) Average information gain: in the process of operation, the difference between two amounts of information is generally expressed by information gain. In the process of selecting classification attributes, the larger information gain of gyroscope is generally regarded as classification attributes.

## 7 My learnings

I learned about the cloud computing and its various implementation in areas like Cloud computing , Artificial Intelligence, Data Science and Machine Learning. Also, I studied about the AWS platform get to know about the different algorithms used in this field that can be used for the solution of the project. We get to know the importance of cloud in today's world. We get to study about various algorithm that can be implemented in this project. I learned how to apply them and how its implementation can be done, and which will give the best result on implementation. I learned how to apply them and how its implementation can be done, and which will give the best result on implementation. I get to used node.js and react.js as fronend and backend to create the website and cosmos Db as database for storing the data.

This 6 Week Internship program with upSkill Campus and UniConverge Technologies Pvt. Ltd. was very much helpful. I learned so much about emerging technologies like node.js and react.js and get to work on a real-world project. It was a great experience which will help me get ahead in my career in future. Thanks to upSkill Campus and UCT for giving us this opportunity.

## 8 Future work scope

1. **Feature engineering:** The performance of decision trees can be improved by processing the input data beforehand and creating useful characteristics. You could design elements like time of day, day of the week, holiday, and seasonality indicators for traffic predictions. Additionally, take into account including traffic-related elements as traffic jams, collisions, and construction activities. The decision tree algorithm can better anticipate outcomes by capturing complicated relationships through feature engineering.
2. **Incorporating real-time data:** Decision trees can be trained using historical traffic data, but real-time data, such as traffic flow, weather, and events, may increase predicting accuracy. The decision tree may adjust and create predictions based on the present circumstances by incorporating these variables, producing more precise traffic forecasts.
3. **Hierarchical decision trees:** Traffic patterns, like daily, weekly, and monthly trends, might show hierarchical systems. You can identify these patterns on several levels by building a hierarchy of decision trees. For instance, while lower-level trees can concentrate on weekly or monthly patterns, the top-level decision tree can forecast daily traffic trends. This strategy can increase forecasting precision across a range of time spans.
4. **Spatial dependencies:** Road networks, traffic congestion, and transportation infrastructure are examples of spatial elements that frequently affect traffic patterns. Consider using geographic information system (GIS) data to incorporate geographical dependencies into the decision tree algorithm. The decision tree can more accurately depict

localized traffic patterns by using spatial variables like road connection, distance to important destinations, and population density.

5. **Dynamic updating:** Various variables, like urban development, modifications to the road system, or significant events, can cause changes in traffic patterns throughout time. Consider adding a dynamic updating mechanism to the decision tree model to take these changes into account. Retrain the decision tree on a regular basis with the most recent data to account for changing traffic patterns and guarantee forecast accuracy.
6. **Ensemble methods:** Accuracy can be raised by using decision tree ensembles like gradient boosting and random forests. Ensemble approaches lessen overfitting and produce more reliable traffic forecasts by training many decision trees and pooling their predictions. Individual decision trees cannot manage noise and oscillations in the data as well as ensemble models, which leads to more accurate predictions.