Abstract

In this Cenozoic era, graduating from University is like a standard procedure for getting a diligent job in the society. University also getting more in numbers and became common for students to achieve their high level studies in every industry. However, informations from every university's website are rarely standardize and some even fragmented which had caused the third party such as students and parents had their hard time to do researches with every single university. This phenomena will bring students into an unfavorable situation where they might choose the wrong university for themselves due to the difficulties in investigating the aspects and beneficial in each university.

This project is aimed in developing for educators in Universities to convert their areas of expertise into prototypes, curriculum, industry-friendly collaboration models and develop new areas of research with AI-enabled engine. This project will discover each educator and his/her specialisation and achievement, then create a personalised AI expert system.

Table of Content

Abstract

Chapter 1 Introduction

- 1.1 Background
- 1.2 Problem Statement
- 1.3 The Company Background of Industry Collaborator
- 1.4 Project Background
  - 1.4.1 Problem
  - 1.4.2 Target Market
  - 1.4.3 Competition/ Contribution
- 1.5 Objectives
  - 1.5.1 Project related Questions
  - 1.5.2 Research Aims and sub-Objectives
- 1.6 Project Scopes
  - 1.6.1 Dataset
  - 1.6.2 Programming Language
- 1.7 Project Plan
  - 1.7.1 Phase 1 Research and Prototype
  - 1.7.2 Phase 2 Website Scraping Module
  - 1.7.3 Phase 3 Data Segmentation Module
  - 1.7.4 Phase 4 Dashboard Module
- 1.8 Project Organisation and Timeline
- 1.9 Advantages and Contributions

Introduction

This project is involved in big data analytic to help universities and industries in creating a linked connection between Universities, Industries and Investors. With the connections, the published researches and uncommercialized patents from the universities able to seen by the investors and industries to turn these universities works become commercialised in the community. The data sets used by this projects involves the details of Universities, the governments, and et cetera.

This project includes functions such as data crawling from various websites, segmenting raw data into data clusters and insert into an online database. These functions are handling by 2 groups which involve with 7 different teammates.

1.1 Background

MyFinB (MFB) is one of the award-winning AI company that frequent plays a role of an AI-as-a-Service (AIaaS) platform for various industry. They focus on building new proprietary analysis platform with excellent tools that able to transform structures and unstructured data into a readable and user-friendly format for the third party.

1.2 Problem Statement

With these loads of universities in Malaysia, it takes time for third party such as the teenagers, parents and lecturers to explore and do researches of the universities with every aspects such as the Courses, Subject material, Social activities, Portfolios and et cetera. However, it is found guilty that the information that we can find in websites of most universities are counted as fragmented information which are all broken pieces of data that are not close together. Some universities even irresponsible to update their university information to the latest such as the alumni of the university and facilities developments. Therefore, these fragmented data became the core problem of third party having difficulties in getting true real-time information.

This core problem, the fragmented data issues in every university had effected two main phenomena in the society. The phenomena could easily discovered are the third party had hard time to produce a better choice in targeting the most suitable university for every individual. Feedbacks of every subject from every student will always collect from the universities after every semester. However, the feedbacks are

disclosed by every universities which actually the feedbacks are very vital to the third party as they able to feel the true learning environment from the feedbacks of the students. Websites and advertising will always show the best from their facilities but feedbacks will show the true feeling towards the surrounding. Thus, It is possible for students to remain anonymous and providing feedbacks in the surveys or forums to show the true experience received from the universities

The second phenomena is the connection between the industry, universities and investors is weak which happens to categorized as part of the fragmented data. It is to believe that every universities have numerous of published researches and non-commercialized patents that need more engagement and adoption by the industries. Moreover, technology is getting more contact with the methods in teachings of educators which change the norms in students learning experience. Hence, industries needs to have a strong connection with the educators to ensure the enhanced learning scope and contents to be learnt by the students in universities.

1.3 Company Background of Industry Collaborator

Figure 1.3.1 UIP, CE.A.I., and MyFinB Logo

MyFinB is a risk analytic platform that uses artificial intelligence to process assessments and insights to the users and industries. With the big need of Artificial Intelligence in next few years, they start providing a platform by gathering all different patents and researches for a efficient way to community and investors to understand and look through all educations data.

1.4 Project Background

The characteristics of this project is based on analysing a huge amount of various universities data that look meaningless and turn them into meaningful by segmentation which able to assist in findings insights from the informations of universities.

1.4.1 Nature of Business

1.4.2 Problem Statement

With these loads of universities in Malaysia, it takes time for third party such as the teenagers, parents and lecturers to explore and do researches of the universities with every aspects such as the Courses, Subject material, Social activities, Portfolios and et cetera. However, it is found guilty that the information that we can find in websites of most universities are counted as fragmented information which are all broken pieces of data that are not close together. Some universities even irresponsible to update their university information to the latest such as the alumni of the university and facilities developments. Therefore, these fragmented data became the core problem of third party having difficulties in getting true real-time information.

This core problem, the fragmented data issues in every university had effected two main phenomena in the society. The phenomena could easily discovered are the third party had hard time to produce a better choice in targeting the most suitable university for every individual. Feedbacks of every subject from every student will always collect from the universities after every semester. However, the feedbacks are disclosed by every universities which actually the feedbacks are very vital to the third party as they able to feel the true learning environment from the feedbacks of the students. Websites and advertising will always show the best from their facilities but feedbacks will show the true feeling towards the surrounding. Thus, It is possible for students to remain anonymous and providing feedbacks in the surveys or forums to show the true experience received from the universities

The second phenomena is the connection between the industry, universities and investors is weak which happens to categorized as part of the fragmented data. It is to believe that every universities have numerous of published researches and non-commercialized patents that need more engagement and adoption by the industries. Moreover, technology is getting more contact with the methods in teachings of educators which change the norms in students learning experience. Hence, industries needs to have a strong connection with the educators to ensure the enhanced learning scope and contents to be learnt by the students in universities.

1.4.3 Target Market

The target market of this project are the Universities researchers, the academic community, Centres of Excellence of Universities, Government agencies and Ministries, SMEs and Corporation in 15 industry Groups, Investors, and Undergraduates and Postgraduate students.

1. University researchers and the academic community

They plays a role of contributing and sharing their research or patent ideas to the platform as the platform tends to accept as much as possible of researches and patents to display and commercialise the education materials to the public

2. Centres of Excellence of Universities

Centres of Excellence in Universities are targeted as they could assist the project outside of the platform to filter out the best team of experts to help out with the researches or the uncommercialised patents.

3. Government agencies and Ministries

Government are one of the target market as we aim to get the attention of the government so that the agencies and ministries willing to plan it as part of their project for the future industries and education institutions for a better wisdom level in the country.

4. SMEs and Corporations in 15 industry groups

Small Medium Enterprise plays a role of invest into various patents to be part of the contributors and use the patent technology or ideas in the community for the increment of their business and economics.

5. Investors (venture capitalists and private equity groups)

We need investors as the platform are a long lasting project and platform for the education community. Cannot be denied that minority of the researches or patents need long lasting equity to continue and succeed in their researches. Hence, Investors from venture capitalists and private equity groups are one of the target market we would capture.

6. Undergraduates and Postgraduate students

We would like undergraduates and post-graduate students to be part of the platform as they could easily find their interest researches from the platform and build up their passion from it. They also able to analyse the universities from the activities and information from the platform and make a good decision to pursue their studies in a right university institutions.

1.4.4 Competition / Contribution

Currently, similar purpose and characteristic platform does not exist yet in Malaysia which are dedicated for academic researcher to have a collaboration with various industries. However, other comparable aspects that are unforeseen would be the accelerators and incubators that are catering more when the start-ups phase.

The comparison will be used during competing will be the difference of UIPFuture platform and other accelerators and incubators platform. UIPFuture platform also will be juxtaposed with Crowdsourcing platform as they plays a role of inviting expertise into various projects.

The UIPFuture platform is purely contributing for academic researchers and lecturers in every Universities that are collaborating with.

1.5 Objectives

The Objectives of the project is obvious as below:

a. To crawl various information from websites into excel
b. To create a software which act as a centralized database to store every information related to the universities.
c. To extract all data from excel into a centralized online database
d. To provide a platform to gather all informations of universities
e. To provide a convenient, efficient and accurate visualisation results to the third party.

1.5.1 Project Related Questions

1) What is the quantity and the quality of the sources can be collected from every university?
2) What will be the main focus of the students when making decision?
3) Will the platform be accurate for the academic researchers when making decision?
4) How to retrieve information from institutions that insist in disclosing their information?
5) How fast will the system be respond and finish analyse the results?

1.5.2 Project Aims and Sub-Objectives

The project aims is to provide a convenient platform for academic researchers' universities researches and uncommercialised patents bring into the real life and various industries to be commercialized.

The project sub-objectives are as below:

a. To improve the skills of webscraping websites and extract from codes instead of displays.
b. To bring up the work and contributions of educators' researches.
c. To simplify the workflow and figure out the effectiveness of the platform.

1.6 Project Scopes

In this project, UIPFuture will be developed and during the development phase, several works and scope in this research will be focused as follow:

1.6.1 Dataset

The dataset will be used in this project are the set of data of every important information in every higher level study institutions, industries and government agencies.

1.6.2 Programming Language

The programming language that we will be using are HTML, CSS, JavaScript, Python, Php and Laravel. HTML, CSS and JavaScript will be used to assist in building user interface; Php and Laravel will be used to build the database; Python will be used to scrap data and analysis.

1.7 Project Plan

UIP, an AI-as-a-Service (AIaaS) platform will be built for educators in Universities to convert and commit their areas of expertise into prototypes, curriculum, industry-friendly collaboration models and develop the new areas of research with AI-enabled engine.

UIP will be created and designed to showcase every select UIP researches with various categories into one single platform. All researches will be categorised as Agriculture, Board Reporting, Business and et cetera. In the platform, user will able to explore the UIP projects based on their interest.

The project also will be undergo with 4 phases as below:

1.7.1 Phase 1 Research and Prototypes

Various researches are selected and prepare to insert into the platform to have a chance of joint collaboration with industries and develop a research prototype from the research itself. The researchers will have the opportunity to co-own the prototype when it is developed.

1.7.2 Phase 2 Website Scrapping Module

During this phase, we will scrap all information we need from the Universities, Government Agency and Industries. It is needed to understand more the connection between three of them to build a highly-usable platform.

1.7.3 Phase 3 Data Segmentation Module

In the phase, we will be gathering all data we got from phase 2 while we create a centralized online database to insert all the data into it. The raw data we collected from phase 2 will also convert into various sets of data clusters to become meaningful for us in the project.

1.7.4 Phase 4 Dashboard Module

At this phase, we will create a dashboard to export the data that are able to assist the user. Before viewed by the specific user, the data will be analyzed by various methods to display out meaningful insights in different formats such as bar graph, top ranking chart and et cetera.

1.8 Project Organisation and Timeline

Timeline	Mile Stone		Mile Stone Goals
16/03/2021 - 30/03/2021	Web Scraping	Universities Staff Module	Scrap all the related information and produce output in .csv format.
25/05/2021 - 29/05/2021		Government Agencies Module	
19/04/2021 - 30/04/2021	UI/UX development	UIPFuture website Module	Create a website as a platform to display various researches
		UIP projects Module	Gather and Categorised researches which collaborated with MyFinB
	Database development	Push all the scraped information into one central database with auto search and filter functions.	
	Database UI/UX development	Development of UI/UX for the database to produce	

		information needed and data flowchart.	
	Dashboard Design and Development	Design and development of the dashboard of the RoboAdvisor AI-based system.	
	Live feeds scraping	News scraping for latest development/trend	

1.9 Advantages and Contributions

The finding of this research study will be reflected to the community and beneficial for the education community such as the universities, industries and the investors as this research will be seeking linked relationship and collaboration towards each of the universities. Academic researchers will also find easy and convenient to do their own finding as well as fund can be easily funded when the researches is easily found by investors in a platform.