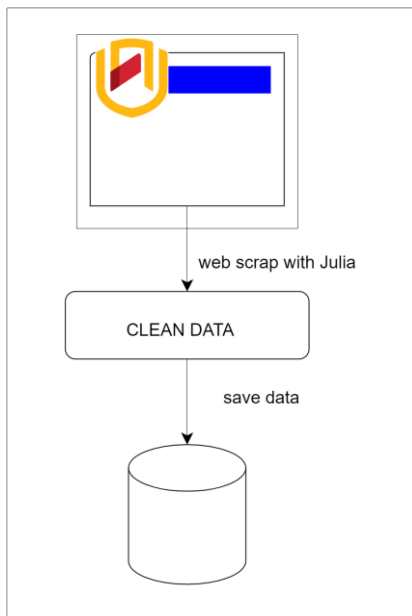


Introduction

In an age that is moving towards Artificial Intelligence, it is important to incorporate it in our businesses and lives. For an institution like the Namibia University of Science and Technology (NUST) it is as important. The powers and capabilities of Artificial Intelligence and Machine Learning can make a user's experience of visiting the NUST website much richer through a Chabot. In this literature review, we look at the approaches which will yield the proposed Conversational Agent for NUST.

The data

In order for the chat bot to answers its users questions it needs knowledge. Thus acquiring the chat bot's data is the first crucial step.



There are plenty of pre collected datasets that could be obtain, but none solely on NUST. It takes a long time to obtain sufficient dataset for a cha bot's intelligence, sufficient being a total of 500 thousand records at least. Most datasets are build based upon for example reddit answers of a score greater than 2, which means that a bot build upon such a dataset will be able to answer almost anything random question, and will some will have a lot of opinions some possibly incorrect and outdated answers.

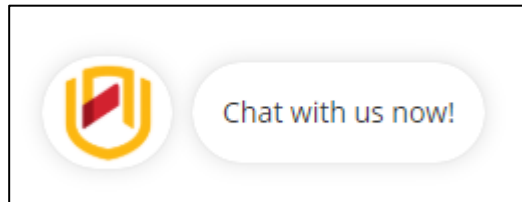
This research's chat bot aim is to have it answer any question on NUST. To combat the curve of having to collect this data, and not have outdated data is to collect the data right off the official NUST website. Julia a high-level, high-performance, dynamic programming open source language provides packages to will scrap the NUST website with ease. It is a well-suited language for high-performance numerical analysis and computational science and will be able to scrap off the website as well as clean it.

The clean data will have an additional dataset that will add some personality to the chat bot, to make it relatable to the youth, and make it more fun to use.

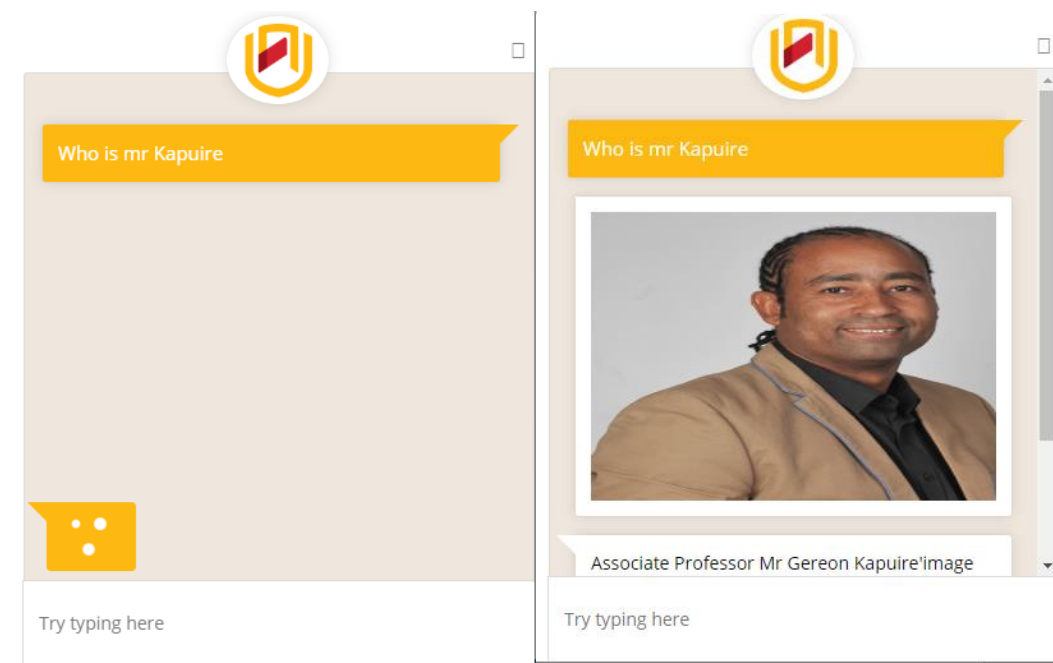
Chat bot Interface

The chat bot's design possibilities are endless and is of the least priorities of this research, but sticking to brand with the NUST website would be most professional and appropriate. Below is the prototype on the chat bot interface on hand currently.

How the chat bot will hover on top of the NUST website



On click, the chatbot opens



The prototype above has the lowest level of intelligence, it in fact has not machine learning models implemented yet and only identifies the noun in the question asked and attempts to find it on the NUST website, which it successfully did in this case with ease by Julia. With machine learning models this prototype will be more efficient.

Machine Learning Model

Machine Learning provides a plethora of ways one can create a chat bot model. A model is the chat bot's brain per say. The clean dataset from the NUST website will be used to train with a model, and the bot will get better and better and answering our questions.

TensorFlow is an end-to-end open source platform for machine learning. Since TensorFlow is originally developed by researchers and engineers working on the Google Brain team within Google's Machine Intelligence Research organization to conduct machine learning and deep neural networks research, it will be a great machine learning tool for this research. It will be used to build the chat bot's model.

TensorFlow on its own is not sufficient, packages like pandas, nltk from python and word tokenizer will be used to aid the building and training of the model.

Deployment

Docker is a set of platform that allows one to containerize an environment. Docker will be used to keep the environment on which the chat bot is deployed constant. A Dockerfile file will be written, to which a docker image will be built from a base TensorFlow image.

Docker will provide an already installed TensorFlow environment which is rather hard to install on one's own computer. The docker image of our chat bot can be pushed to Docker hub, for safety and backup.

Conclusion

The reviewed literature suggests that there the experience of a NUST website user will be greatly improved by having a conversational agent to interact with the user. Directly getting the data from the website will assure the chat bot is always accurate and reliable. User's will get the impression of almost talking to a person, with a brain, a well-trained TensorFlow model that is. The final product will surely attracted a lot of people to the NUST website.

References

1. Aurelien, G. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts (2nd ed.). Sebastopol, CA, NJ: O'Reilly Media.
2. Babak, R., Harrison, B., Mohammad, A. (2017). An Introduction to Docker and Analysis of its Performance Retrieved from https://www.researchgate.net/publication/318816158_An_Introduction_to_Docker_and_Analysis_of_its_Performance
3. Giancarlo, Z. (2016). Getting Started with TensorFlow (Kindle Edition). Sebastopol, CA, NJ: Packt Publishing Ltd.
4. Kinsley, S. (2020). Python Programming / NLTK-3----Natural-Language-Processing-with-Python-series. Retrieved from <https://github.com/PythonProgramming/NLTK-3----Natural-Language-Processing-with-Python-series>