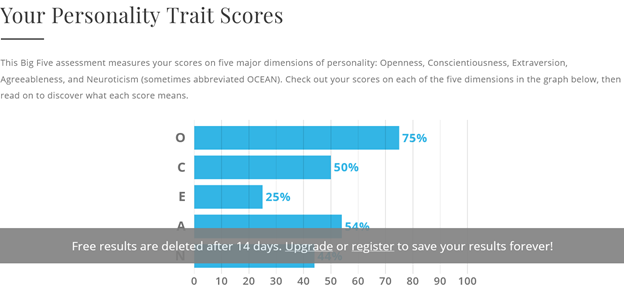
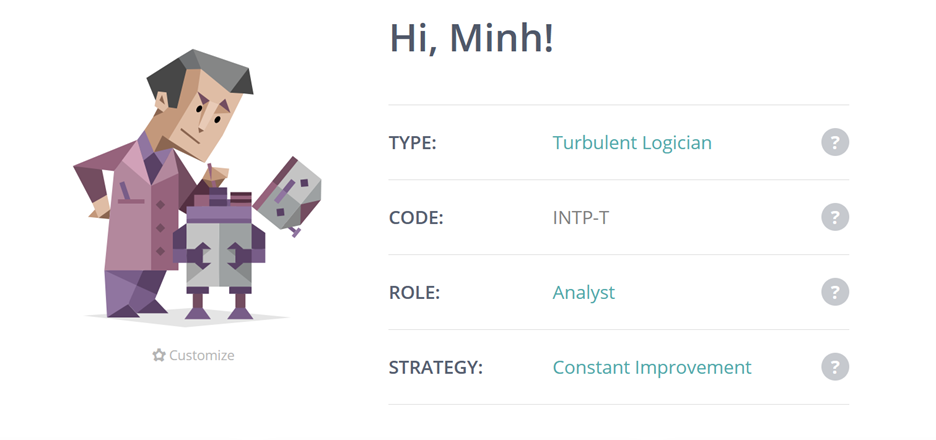
# Team Profile

1. Team name:
   * I have decided to name my team” Tekno”. I chose this name as this seems to be futuristic and foreign and is seems to be related to technology.
   * Team ProfileGraphical user interface, text, application, email

     Description automatically generated





Here, I have included the results of 3 different online tests that I have completed. In the presented order, these are the learning style test, the Big-5 personality trait test, and the Myers-Briggs test. I don’t often choose my teammates based on the results of personality tests. This is due to the fact that some test such as the Myers-Briggs test have been criticized greatly for its lack of reliability and its lack of account for some complex personality type. But overall, if were I need to use personality tests, there are 3 factors that I use to choose my potential teammates: how they behave under pressure, how they are willing to accept new ideas, how do they act when they collaborate. So, I will mainly use the big 5 test for my choices as I believe the 5 aspects that the test measures will quantify the traits that I want my team mates to have. To find someone who is good at dealing with stress, I will look into the Neurotic quantity as it identify how one is going to react in stress. I will also investigate their Agreeableness and their openness quality to satisfy my next two requirements.

# Tools

Website link: <https://duplominh.github.io/COSC2083-ASM2/>

GitHub: <https://github.com/DuploMinh/COSC2083-ASM2>

The GitHub history showed that in the beginning I made a lot of commits with minimal change as I was trying to iron out the configuration of the GitHub page site. Later, as I settled on a configuration, I only added content to the page as I am satisfied with how the website look.

# Project Ideas

## Overview:

My project is about creating a grammatical correction engine with Vietnamese texts as Grammarly has done with English. I wanted to create an engine that can correct grammar mistakes, typos, incorrect word choices and incorrect position of accents in Vietnamese texts. I intended that this service could be delivered over the Internet such as the case with Grammarly or it could be provided as a software package to run locally on a machine such as with the current service offered in Microsoft Word.

## Motivation:

I thought of this idea when I heard a friend complaining that he had to proofread papers as a part of his part-time job. I then started to wonder if there is a way to automate this time-consuming task of checking for grammatical mistakes and typos and make the job of proofreader easier as they now can focus more on the content. This product has potential as even native speaker of a language can make mistakes, therefore, having something that can automatically suggest, and correct language mistake will have a market as Grammarly has proven.

## Description:

If you are a student or rather anyone working often with English documents, you are often familiar with Grammarly. Grammarly is a service that, from documents, tags and suggest corrections to grammar, typos, and word choices. I plan to replicate this service, but instead of doing this with English documents, I plan to do this with Vietnamese documents. This task could be defined as a translation task in Natural language processing. By using the source document as the source language, it could be “translated” into standard Vietnamese. The system should be able to tag the changes has been done to the source document and identify what kind of change is it. Then from the tags, I can derive my features. The features are:

+ Identification of grammatical mistakes.

+ Identification of typo. This includes incorrect placement of accents, mistake in use of “vần” such as “ch”, “tr”

+ Suggestion for more appropriate word choices in the context.

I planned to deploy this service in 2 ways. First, I will have a website that allows user to upload documents in which it could be analyzed, and the user can get the result immediately. Another method that I planned to deploy this application is to release addons for applications such as Google Docs or Microsoft Word. This allowed the user to get correction for their documents as they work on it. An additional method that this app could be deployed that I will investigate is as an API. Users can make an API call with the text that should be analyzed, and the service will return the analyzed text that has the suggestion tags.

## Tools and Technologies**:**

I planned to base my project on the GECToR – Grammatical Error Correction: Tag, Not Rewrite project by Grammarly ([arXiv:2005.12592v2](https://arxiv.org/abs/2005.12592v2) [cs.CL]). By using their framework which runs on PyTorch which is a machine learning platform, I would be able to reduce my workload to adapting the tagging system to the specifics of Vietnamese. I would also have to create a new dataset for training as there is no commonly available dataset to train the models with Vietnamese language features and quirks. For the front end of this product, I planned to use C# to be a cross platform front end for this application. For the training of the models and the delivery of the services, I intended to use cloud service providers to provide the computing power as well as distribution of the product.

## Skills Required:

My project needs an advanced understanding of Natural Language Processing and Machine Learning. I will also need to create a front end for the user to be able to use the product. There should not be any special needs of hardware as cloud computing providers can provide the computing power that I need like high performance computers to train models and high availability servers and content delivery network to deliver the service.

## Outcome:

I don’t intend for this project to solve any great problem of our society. Rather, I oriented my project to save people’s time. Therefore, I am contributing to solving the problem of our society by allowing people to focus more on solving problem rather than spending that time on some mundane tasks such as correcting grammar mistakes. Another outcome of this project is that I am able to satisfy my passion in making a project involving machine learning.

# Feedback

Feedback 1:

“The project proposal seems to have explained well the intention of the project. However, some technical part should have been elaborated more on and some keyword should have been better explained. Also, maybe you should include a business model for your project”

Feedback 2:

“I think overall the idea presented here is fine. However, I question the feasibility of the project as many of the technical aspect seems to be complicated and it is not thoroughly explained how he is going to overcome the technical challenges.”

I think the feedback that I received here are completely valid and constructive. I do think that I could explain the technical terms in my project proposal a bit more. However, I could justify for some of my choices in the proposal. One of the points of the first feedback is that I could have included a business plan in my project. For this, the reason that I have not included a business plan or model for my project is that I believe this is not included in the requirement of the exercise to include a business model or how I am going to monetize this project. Including this should be easy as with how the project is proposed I could include this by selling the access to the service or application. For the technical aspect and how it might be not feasible. For this, I believe that the framework that I have chosen is quite adaptable to different language and it has been done by different teams with minimal adjustment to the underlying model. As I don’t have to modify the model, I will not need a too great understanding of NLP. Therefore, I think that I can execute this project.

# IT technologies

The two areas that are most relevant to my project are Natural language processing and Artificial Intelligence/Machine Learning.

## Natural Language Processing (NLP)

Natural Language Processing is the study of interaction between computer and human language, notably to process natural language data. Natural language processing final goal is for the computers to have an “comprehension” of human languages with its nuances. Comprehension here means that computer can extract the meaning out of texts like a we as humans could. Currently, usage of NLP is widespread and could be seen in every aspect of life. When you wake up, you might ask a virtual assistance for your schedule that day. This task is a combination of Speech recognition, speech synthesis, and Question answering. When you search up a name on Google, for example “John Major”, and the result returned is about John Major the British prime minister instead of a Major named John or a John who majored in some subject, this is called lexical semantics or understanding of individual words in a context. When you drive out of your house and your license plate is recognized at a toll booth, it is Ocular Character Recognition. When you use an online translation service to translate some text, that is machine translation. These are just some of the more notable advances and usage of the field of NLP. There are also some notable advances in NLP that are less publicly known. One of such advancement is in the field of natural language generation. This is notably used to generate textual report from a data set, for example a weather forecast, a business report, or a chat bot. Multiple NLP disciplines has been used in combination to create a Dialogue management system or a system intended to hold conversation with humans. One of its implementations is with Google Voice where this system can make reservation over telephone in place of a human. Currently and in the near future, advancement of NLP is closely tied with artificial intelligence and machine learning as the pattern matching ability that AI and ML offers enable is perfect for a wide amount of NLP task where pattern recognition and matching is the core of the problem. Future advancement therefore will also be tied to AI and ML. One of such advancement will be the speech tonal stress reconstruction. Systems in the future will be able to generate speech that more closely resemble such of human by replicating the stress and syntax that human makes. Cloud computing is another enabler for advancement in natural language processing as the processing power needed in certain NLP application can be provided over the internet. This allowed the application platform for NLP to extend beyond large desktop or mainframe and go to every networked device. This includes virtual assistant which can appear in small device such as smart speakers, smart camera that can read license plate, and many more that utilized cloud computing to put NLP application everywhere.

NLP had and will change our life greatly. In many fields, we can already see NLP in some part being applied in different forms. It could be an application of Ocular character recognition to transcribe text, or it could simply be a spam filter for emails. In the future, we will see even more application of NLP being used daily. NLP has already replaced human jobs. The most notable is with help desk and customer support jobs. These are occupations that is being and is highly likely to be replaced by NLP application such as chatbots and voice assistant. These occupations are the perfect use cases for NLP application. On the other hand, for NLP application to spread, we will need more NLP applications developers.

Personally, I do not think that NLP will affect myself, my family or anyone that I know of in a significant way. Although I and quite a few people I know have been in contact with many applications of NLP, these interactions have not been too influential on my life. I think certainly NLP had, have, and will make much of my everyday life faster. However, NLP has not changed drastically how I lived.

## Artificial Intelligence / Machine Learning:

Artificial intelligence refers to “intelligence” being displayed by machine as opposed to natural intelligence displayed by humans and animals. Machine learning is a subset of AI which studies computer algorithm that can improve on its own through “learning” and the processing of data. Nowadays, as development of the two subjects mentioned here are well interconnected, the two terms AI and ML are often mistakenly used interchangeably but it is to be noted that these are two, while related but different subject. Here, I will discuss mostly about ML and AI in relation with ML since AI is a relatively basic concept that covers a great deal of subject from technical to philosophical and the depth of information is unfathomable. Machine learning can be divided into smaller subjects, each with an interesting state of the art development. For computer vision, Semantic segmentation is one of the notable tasks. Semantic segmentation is the analysis and understanding of an image into structure and component in the pixel level. This technique is used in self driving function for cars to analyze the image of the road to guide the car in conjunction with other sensors. Another notable task for machine learning is image classification. This is, from a given image, the model can classify an image into a category. Object detection is also another notable ML task where given an image, objects inside that image can be identified and classified. Many natural language processing tasks such as machine translation or language modeling are also machine learning tasks. In the future, there are many ways that machine learning can develop. One of such ways is with automated machine learning. Here, a machine learning model automatically select the appropriate algorithm for the dataset. This will allow even non-ML experts to deploy ML models and facilitate the usage of ML more widely. ML will also be applied in more sector than now. Healthcare is a subject where ML can be applied well as the dataset generated in healthcare are often massive and is hard to analyze with other data analysis technique. For example, disease prediction could be done by ML by analyzing a range of variables for each patient such as health history, demographics and other info. Machine learning can also help in the drugs development process by speeding up the testing process using models and help to reduce the costs. Self-driving vehicles is one of the prospective sectors in the future for machine learning. Even though ML is currently used in this sector intensively, problem arises when the models must deal with untrained condition. In the future, new algorithm and method may allow better extrapolation of the models used in self driving. The growth in AI and ML are enabled by the increasingly more powerful processing power in computer. Transistor counts on an average computer chip has increased by five million times from 1980 to now, with the corresponding increase in processing power. This allowed more complex models to be created and operated. Also, the increasing amount of data available that cannot be processed in traditional methods meant that new method must be developed, led to development in ML.

Artificial Intelligence and Machine Learning will change our life greatly, however, we will have to wait for that. AI have the potential to drive for innovation in many sectors. For the near future, I believe that AI and ML will not make redundant too many occupations. This is because, AI can only focus on one task only, one problem at a time, also known as specialized intelligence. Humans have general intelligence, therefore, are more flexible and can perform many tasks. Also, the data to train the AI are not perfect and may be subjected to biases that can affect the AI.

# IT work

Here, I have decided to interview Mr. Phan Trọng Huy. Mr. Huy is the Director of a software company that I have interned in. He is also the project manager so he could be considered an IT professional. My interview with him is as followed.

Question(Q): *Thank you for agreeing to this interview. Can I ask what is your current position?*

Answer(A): You are welcome. I am currently the director of Khang Vân Co LTD. We are specialized in software and specifically shipping management software.

Q: Can you tell me what kind of work do you do day to day?

A: Well, as the director, I am responsible for a great deal of business, especially as we are a small company. I work with customers to find out what do they want the application to achieve and to plan the features that will then be sent to developer and designers to develop. I also work with the customer during the development process to receive their feedbacks and customize our application to satisfy the customer. I also supervise the deployment of the application and helps our client with some of their needs during this process such as setting up the necessary infrastructure. During the development process, I plan how the feature should be developed by managing the timetable and specify the technicalities of the features such as what should an API returns or the arrangement of a database system.

Q: What kind of people do you have to deal with?

A: I deal with all kinds of people. In our company, I work with my employees to coordinate business. I work with customers as I have mentioned before to develop our products. I also work with third parties such as vendors or port authority to secure the necessary cooperation that we need in the development process. For example, we work with port authorities to integrate their system into our product that we then deploy for our customer.

Q: Where do you spend most of your time?

A: My time is often divided between in the office and meeting with customers to discuss the products.

Q: What do you find most challenging about your job?

A: I would say having to cater to customers’ requirement. This is because many of the time customers would request something that is absurd that we must talk them down or they cannot accurately describe what they want so when we deliver them the prototype, they don’t like it, so we have to change it. Customers many of the time don’t know what they want so we must work with them to deliver a product that with satisfy them.