# Introduction.

* 1. Project overview.

The project we are doing is based upon Machine Learning and we are aiming to develop a swahili transcription tool using machine learning. We will develop a model using machine learning natural language processing algorithms and finely tune the model to give the highest percentage of accuracy in transcribing the swahili language words into swahili text.

We will later deploy the model that we will develop into a web based application using a python web application building framework, thus creating an acoustic model and a batch streaming application of the model.

Developing this tool we are expecting to have the first swahili transcription tool created in Tanzania, as swahili language is originally from the bantu people of the eastern Africa decent. With unlike the rest of the transcription tools that have been developed by different developers from universities and colleges around Europe and Australia which have no great knowledge of the swahili language compared to the indigenous people of East Africa.

* 1. Problem Statement.

Artificial Intelligence is preferred as a technology that solves human problems with ease, it uses computer sciences with programming knowledge for computers to understand simple instructions from human intelligence and later mimic the human behavior at a faster and more accurate level (Robert Fay, 2021). The subset of Artificial Intelligence which is Machine Learning (ML) is the leading of this advancement. This advancement has led to the development of tools that assist humans with different tasks such as transcribing and translation, with ease, accuracy and fast execution.

Natural language processing technology in the Artificial Intelligence and Machine Learning computer sciences has grown and developed very greatly in the past couple of years, as seen with its implementation in multiple technological companies through the use of chat bots and Artificial Intelligence assistance technologies. For example; Apple having Siri and Samsung having Bixby. All of these have been made through the use of natural language processing technology in machine learning. (Ma, 2022).

Focusing on transcribing, Machine Learning has enabled people in need of this service to finally be able to acquire the resources and services at a much cheaper price and with faster

Execution of the service with little time consumption (Ranchal, 2013). Transcription tools are available in multiple language, even swahili, but with the swahili language there is a low accuracy of the AI model. We have realized that AI models, the natural language processing field struggles with accents even with the English language and so the AI models lack a bit of accuracy due to accents (Hyryn, 2020), then the swahili language also has multiple accents that vary with the indigenous people within Tanzania. This is may be cause of the fact that the data for the manipulation of the model did not originate from the swahili people parse. The existing models fail even to capture simple swahili words like “ujenzi”, “maisha”, “uzoefu” and lots of more other words that can frequently be heard from the swahili people especially in the swahili speaking regions in Tanzania like Dar Es Salaam (Masua, 2020).

With the Swahili transcription tool, we will have the first transcription tool built by Tanzanians and with the implementation of speech recognition optimization a normal Swahili accent will be used as a benchmark for the rest of accents to fall through, thus increasing the accuracy of this tool far greater than the ones built by Swahili scholars abroad who are not well acquainted with the Swahili language than the indigenous people of Tanzania where Swahili is largely used compared to other parts of the world.

* 1. Project Objectives.

In this section we will be evaluating the objectives that we have set for the accomplishment of our project, the goals that we are meaning to attain at the end of the entire project.

* + 1. General Objectives
* To develop a swahili transcription tool using machine learning.
  + 1. Specific Objectives
* To find, download and prepare a swahili words dataset for transcription.
* To develop a swahili transcription model using natural language processing algorithms.
* To implement the model by deploying it in a web-based application.

# Literature Review.

There are some transcription tools that we have come across in the search for previous works done by developers based upon Machine Learning and AI, these tools are used for transcription and translation. Platforms such as happy scribe from Barcelona <https://www.happyscribe.com/transcribe-swahili> of which offers transcription services for the swahili language, it still hasn’t reached a desired accuracy for its transcription tool, The model is estimated to have an 85% accuracy in transcribing but it is still inaccurate with the Swahili language as the people who developed it are not originally swahili speakers.

There are other tools for transcription such as vocalmatic <https://vocalmatic.com/languages/transcribe-tanzanian-swahili-to-text> from Toronto, Ontario in the united states of America. This tool offers transcription services but it has a low accuracy when it comes to the swahili language, the problem it encounters is still the same with the happy scribe tool which is originally being unable to recognize the swahili speech in its original accent from the roots of the language. Natural Language Processing (NLP) always struggles with understanding human language when it is spoken with a different unfamiliar accent, not only in swahili but also other languages’ as well.

When it comes to swahili transcription tools and their inaccuracy mainly due to the fact that swahili text and voice datasets are very inadequately found. The swahili language is under resourced and the amount of data available for transcription that has been used previously in creating transcription tools is inaccurate compared to other languages (Hadrien Gelas, 2012).

The accuracy of any machine learning can be determined by the amount of data the model is fed to interpret and understand human instructions giving the most accurate output and predictions of the model. The data itself can not only be used for transcription alone but also translation and as computers are to be programmed to mimic human behavior and so the same challenges, obstacles that humans undergo while performing the transcription tasks should be taken into account for, The data collection, study, analysis and documentation should further more consider the fact that languages’ vary between speakers and also the clarity of audio files can have an effect upon the understanding of the language for both human level understanding and machine level of understanding (Himmelmann, 2018).

# Methodology.

# Plan and Budgeting

# References.

# Appendices.