

PROFESSIONAL TRAINING REPORT
at
Sathyabama Institute of Science and Technology
(Deemed to be University)

Submitted in partial fulfillment of the requirements for the award
of Bachelor of Technology Degree in Information Technology

By
Immaculate Maria Fernando

38120031



DEPARTMENT OF INFORMATION TECHNOLOGY
SCHOOL OF COMPUTING
SATHYABAMA INSTITUTE OF SCIENCE AND
TECHNOLOGY
JEPPIAAR NAGAR, RAJIV GANDHI SALAI,
CHENNAI – 600119. TAMILNADU

JANUARY 2021



SCHOOL OF COMPUTING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the Bonafide work of who carried out the project entitled "Speech-to-Text and Language Translator" under our supervision from August 2020 to December 2020.

Internal Guide

Mr. Saravanan P. M.E.,

Head of the Department

Dr. R. Subhashini M.E., Ph.D.,

Submitted for Viva voce Examination held on _____

Internal Examiner

External Examiner

DECLARATION

I hereby declare that the Project Report entitled "Speech-to-Text and Language Translator" done by me under the guidance of Mr.Saravanan P. M.E., at Sathyabama Institute of Science and Technology (Deemed to be University), Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai - 600119 is submitted in partial fulfillment of the requirements for the award of Bachelor of Technology degree in Information Technology.

DATE: 23/01/2021

Place: Chennai

SIGNATURE OF THE CANDIDATE

Immaculate Maria Fernando

ACKNOWLEDGEMENT

I am pleased to acknowledge our sincere thanks to Board of management of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey our thanks to **Dr. T. SASIKALA M.E., Ph.D., Dean, School of Computing** and **Dr. R. SUBHASHINI M.E., Ph.D., Head of the Department, Department of Information Technology** for providing us the necessary support and details at the right time during the progressive reviews.

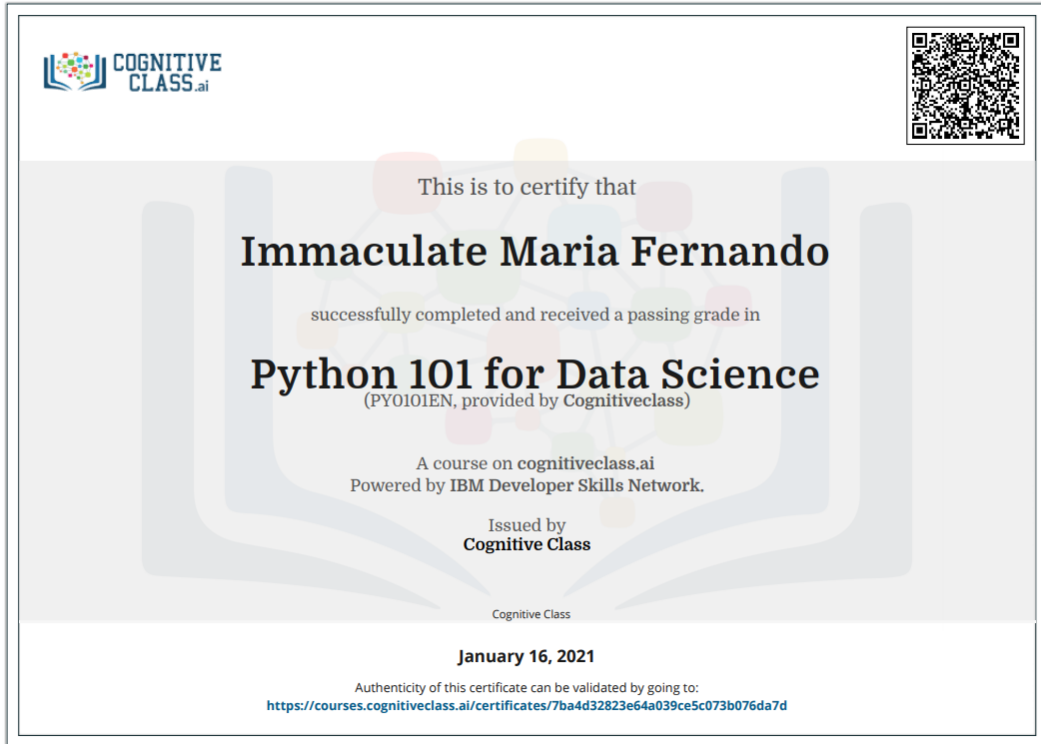
I would like to express our sincere and deep sense of gratitude to our Project **Mr. Saravanan P. M.E.**, for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express our thanks to all Teaching and Non-teaching staff members of the Department of **INFORMATION TECHNOLOGY** who were helpful in many ways for the completion of the project.

TRAINING CERTIFICATE

1/17/2021

Cognitiveclass PY0101EN Certificate | Cognitive Class



<https://courses.cognitiveclass.ai/certificates/7ba4d32823e64a039ce5c073b076da7d>

1/2



[Additional Details](#)

Python for Data Science

Issued by [IBM](#)

The badge earner is able to write their own Python scripts and perform basic hands-on data analysis using IBM's Jupyter-based lab environment.

Skills

Data Science

Pandas

Python

Python Programming

Earning Criteria



[Successful completion of the course "Python for Data Science".](#)

ABSTRACT

SPEECH-TO-TEXT AND LANGUAGE TRANSLATOR

There are many languages being spoken in the world. Not everyone of us is able to communicate through the language of each other. For easy communication, there requires an intermediary between people speaking different language. The Speech-to-Text API takes in the user's voice input and transcribes it. This text is then sent through the Language Translator library to process and we can specify the language to be converted to and the output is displayed. This can be embedded in various sites and app listings as the Language Translator and Speech-to-Text Converter.

CHAPTER NO	TITLE NAME	PAGE NO
	ABSTRACT	6
	LIST OF FIGURES	8
	LIST OF ABBREVIATION	9
1	INTRODUCTION	10
	1.1 GENERAL	10
	1.2 OUTLINE OF THE PROJECT	10
2	AIM AND SCOPE	11
	2.1 AIM	11
	2.2 PROBLEM STATEMENT	11
	2.3 SCOPE	11
3	SYSTEM ANALYSIS AND DESIGN	11
	3.1 GENERAL	11
	3.2 SOFTWARES USED	12
	3.3 ALGORITHM	13
	3.4 PROJECT DESCRIPTION	13
4	RESULT AND DISCUSSION	14
	4.1 SYSTEM REQUIREMENTS	14
	4.2 RESULT	14
5	CONCLUSION AND FUTURE WORK	14
	5.1 CONCLUSION	14
	5.2 FUTURE WORK	14
	5.3 REFERENCE	15
	APPENDIX	15
	A.SCREENSHOTS	15
	B.SOURCE CODE	17

FIGURE NO	FIGURE NAME	PAGE NO
.1	OUTPUT	15

LIST OF ABBREVIATIONS

ABBREVIATIONS

py
ipynb

EXPANSION

Python
Python Notebook

CHAPTER 1

INTRODUCTION

1.1 GENERAL

Using a Speech-to-Text and Language Translator can be very helpful when considering the amount of applications benefited by it. Transcribing text has always been done manually. Using the IBM Speech-to-Text API, any voice message or audio can be easily transcribed. The Language Translator takes in a string of characters as input and converts it into the language one desires.

1.2 OUTLINE OF THE PROJECT

This project is the base for the embedding part in websites. With this has a background, websites can now be able to convert any webpage into the language the user specifies without deviation from the meaning and context. Speech-to-Text can be used in many voice typing like the ones that Google and YouTube support. With further development, this can be scaled to many applications.

CHAPTER 2

AIM AND SCOPE

2.1 AIM

The main objective of this project is to use the Speech-to-Text and Language Translator library to convert an audio file to text and translate into the mentioned language

2.2 PROBLEM STATEMENT

To solve the barriers in communication and make the process more easier and simpler with the use of minimalistic technologies.

2.3 SCOPE

The goal of this project is to be able to embed the notebook into any form of website or app to make the transcribing and translating part more simpler. This project utilizes the libraries provided by IBM to produce the expected outcome

CHAPTER 3

SYSTEM ANALYSIS

3.1 GENERAL

3.1.1 IDE (INTEGRATED DEVELOPMENT ENVIRONMENT)

IDE stands for Integrated Development Environment. It's a coding tool which allows you to write, test, and debug your code in an easier way, as they typically offer code completion or code insight by highlighting, resource management, debugging tools.

3.1.2 USER INTERFACE

The user interface (UI) is the point of human-computer interaction and communication in a device. This can include display screens, keyboards, a mouse and the appearance of a desktop. It is also the way through which a user interacts with an application or a website.

3.2 SOFTWARES USED

3.2.1 CODE EDITORS USED

JUPYTER NOTEBOOK(ANACONDA)

The Jupyter Notebook is the code editor used to implement this project. It provides an easy User Interface and is used to run the code and display it in the same block. Various libraries and methods are invoked and handling huge amount of data is made simpler by using the features provided by Anaconda in its Jupyter Notebook.

3.2.2 PROGRAMMING LANGUAGES USED

PYTHON

Python is a high level programming languages with a large set of libraries and frameworks. Due to its simplicity and the reduction in the amount of code lines, Python has found its applications in many domains. The large set of libraries available in Python is used for Scientific Computing, Data Manipulation and so on.

3.2.3 LIBRARIES AND FRAMEWORKS USED

Pandas

Pandas is a data manipulation tool provided by Python. In the context of this project, Pandas framework will be used for manipulating the data tables and displaying a clear output that is more structured and easy for the users to understand.

3.3 ALGORITHM

IBM Watson

IBM Watson is one of the packages that is provided by IBM containing a huge set of libraries that perform various tasks. In this project two libraries : Speech-to-Text and Language Translator libraries are used for getting the tasks done. As the name suggests, the first library transcribes the text and the second library converts it into the needed language.

3.4 PROJECT DESCRIPTION

This project is the baseline for embedded translators and transcribers in websites. This can help the visually impaired by transcribing a web page and also, the website can display any language the user specifies without changing the context and the meaning of the project.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 SYSTEM REQUIREMENTS

- **Hardware**
 - Laptop/Desktop
 - RAM – 4GB
 - Stable Internet Connection
 - Mic or Audio Recording Facilities
- **Software**
 - Kaggle/Jupyter Notebook
 - API key and URL

4.2 RESULT

A Python Notebook with easy transcribing and language translating abilities.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 CONCLUSION

This project is implemented in the Jupyter Notebook kernel by Anaconda.

5.2 FUTURE WORK

This project is made on the minimalism rule which shows the productive nature of user. It can improve by adding more features to it like timer and clock integration.

5.3 REFERENCES

1. <https://www.ibm.com/cloud/watson-studio>
2. <https://python.org>

APPENDIX

A.SCREENSHOTS WITH SOURCE CODES

```
File Edit View Insert Cell Kernel Help Not Trusted | Python 3.7

In [1]: pip install ibm_watson wget

Collecting ibm_watson
  Downloading ibm-watson-5.1.0.tar.gz (382 kB)
    382 kB 9.6 MB/s eta 0:00:01
Collecting wget
  Downloading wget-3.2.zip (10 kB)
Requirement already satisfied: requests<3.0,>=2.0 in /opt/conda/envs/Python-3.7-main/lib/python3.7/site-packages (from ibm_watson) (2.24.0)
Requirement already satisfied: python_dateutil<=2.5.3 in /opt/conda/envs/Python-3.7-main/lib/python3.7/site-packages (from ibm_watson) (2.8.1)
Collecting websocket-client<=0.48.0
  Downloading websocket_client-0.48.0-py2.py3-none-any.whl (198 kB)
    198 kB 37.9 MB/s eta 0:00:01
Collecting ibm_cloud_sdk_core<=3.3.6
  Downloading ibm-cloud-sdk-core-3.3.6.tar.gz (35 kB)
Requirement already satisfied: urllib3<=1.25.0,!1.25.1,!1.25.1.1 in /opt/conda/envs/Python-3.7-main/lib/python3.7/site-packages (from ibm_cloud_sdk_core) (1.25.9)
```

```
File Edit View Insert Cell Kernel Help Not Trusted | Python 3.7



## Import SpeechToText API



Now, we can import SpeechToText library from ibm_watson. To access an API, we require two things:



1. Endpoint URL
2. API KEY



These two specifications are solely meant for an individual as each person has their own key and url. To authenticate with the cloud server, we require an Authenticator which is imported from IBM cloud services.



In [3]: from ibm_watson import SpeechToTextV1
import json #handling the json file
from ibm_cloud_sdk_core.authenticators import IAMAuthenticator #api key authentication



## Entering the Endpoint URL and API key



In [5]: url_s2t = "https://api.eu-gb.speech-to-text.watson.cloud.ibm.com/instances/7e045591-8076-4362-a02d-ecee6a929504"
api_key_s2t = "9gGYJqYbxogQOCZfDP0lG8gnCi8b32qDirm2Ku0k-DR"
```

```
File Edit View Insert Cell Kernel Help Not Trusted | Python 3.7



## Creating SpeechToText Adaptor Object



In [7]: authenticator = IAMAuthenticator(api_key_s2t)
s2t = SpeechToTextV1(authenticator = authenticator)
s2t.set_service_url(url_s2t)
s2t

Out[7]: <ibm_watson.speech_to_text_v1_adapter.SpeechToTextV1Adapter at 0x7f71f813eb10>



## Download the audio file



In [8]: !wget -O PolynomialRegressionandPipelines.mp3 https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/PolynomialRegressionandPipelines.mp3
--2021-01-22 13:44:59-- https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/PY0101EN/labs/PolynomialRegressionandPipelines.mp3
Resolving s3-api.us-geo.objectstorage.softlayer.net (s3-api.us-geo.objectstorage.softlayer.net)... 67.228.254.196
Connecting to s3-api.us-geo.objectstorage.softlayer.net (s3-api.us-geo.objectstorage.softlayer.net)|67.228.254.196|:443... connected.
HTTP request sent, awaiting response... 200 OK
```

FileEditViewInsertCellKernelHelp

Not Trusted | Python 3.7

In [8]:
filepath = "PolynomialRegressionandPipelines.mp3"

Store the file path

In [9]:
`with open(filepath,mode='rb') as wav:`
response = s2t.recognize(audio=wav,content_type="audio/mp3")

Passing the audio file to the SpeechToText API

In [10]: response

Out[10]: <ibm_cloud_sdk_core.detailed_response.DetailedResponse at 0x7f00108aaaf50>

```
File Edit View Insert Cell Kernel Help Not Trusted | Python 3.7

In [9]: from ibm_watson import LanguageTranslatorV3

Enter the url and api key

In [10]: url_lt = "https://api.eu-gb.language-translator.watson.cloud.ibm.com/instances/e131cb23-8b2a-4962-8433-ba03208fda0a"
api_key_lt = "3pLdZXQed4Qw3M1VNP2rPB_OTz5Vvnwhi7AC_c_2DqBs"

Enter the version of the Language Translator

In [11]: version_lt = "2018-05-01"
```

```
Create Language Translator Object

In [24]: authenticator = IAMAuthenticator(api_key_lt)
lt = LanguageTranslatorV3(version = version_lt, authenticator = authenticator)
lt.set_service_url(url_lt)
lt

Out[24]: <ibm_watson.language_translator_v3.LanguageTranslatorV3 at 0x7f000104df50>
```

```
Edit View Insert Cell Kernel Help Not Trusted | Python 3.7

In [43]: english_translation = lt.translate(text=spanish_translation, model_id="es-en").get_result()

In [44]: new_eng_translation = english_translation['translations'][0]['translation']

In [45]: new_eng_translation

Out[45]: 'in this video we will cover the polynomial regression and the pipes '

In [58]: tamil_translation = lt.translate(text=new_eng_translation, model_id="en-ta").get_result()

In [59]: tamil_text = tamil_translation['translations'][0]['translation']

In [60]: tamil_text

Out[60]: 'இந்த வீடியோக்களில் நாம் போலியோமியியல் ரீதியான மற்றும் பைபைட்களை மறைக்கும். '
```