

## Configuration Manual

MSc Research Project MSc Data Analytics

Abhijith Reddy Pesaru Student ID: x22157131

School of Computing National College of Ireland

Supervisor: Mr Arjun Chikkankod

#### National College of Ireland Project Submission Sheet School of Computing



Student Name:	Abhijith Reddy Pesaru
Student ID:	x22157131
Programme:	MSc Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Mr Arjun Chikkankod
Submission Due Date:	15/12/2023
Project Title:	Configuration Manual
Word Count:	498
Page Count:	7

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

<u>ALL</u> internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Pesaru Abhijith Reddy
Date:	15th December 2023

#### PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	
Attach a Moodle submission receipt of the online project submission, to	
each project (including multiple copies).	
You must ensure that you retain a HARD COPY of the project, both for	
your own reference and in case a project is lost or mislaid. It is not sufficient to keep	
a copy on computer.	

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only		
Signature:		
Date:		
Penalty Applied (if applicable):		

## Configuration Manual

# Abhijith Reddy Pesaru x22157131

## 1 Introduction

This configuration manual walks through all the software and Hardware requirements necessary for this project along with libraries required for building models and all the other essentials will be discussed step-by-step inorder to successfully implement this project.

## 2 Hardware Specifications

#### 2.1 Hardware on local machine

OS Name	Microsoft Windows 11 Home Single Language
Version	10.0.22631 Build 22631
Other OS Description	Not Available
OS Manufacturer	Microsoft Corporation
System Name	LAPTOP-DE1VAHVT
System Manufacturer	HP
System Model	HP Pavilion Gaming Laptop 15-dk2xxx
System Type	x64-based PC
System SKU	48U95PA#ACJ
Processor	11th Gen Intel(R) Core(TM) i5-11300H @ 3.10GHz, 3110 Mhz, 4 Core(s), 8 L
BIOS Version/Date	Insyde F.21, 24-10-2022
SMBIOS Version	3.3
Embedded Controller Version	76.41
BIOS Mode	UEFI
BaseBoard Manufacturer	HP
BaseBoard Product	88E4
BaseBoard Version	76.41
Platform Role	Mobile
Secure Boot State	On
PCR7 Configuration	Elevation Required to View
Windows Directory	C:\WINDOWS
System Directory	C:\WINDOWS\system32
Boot Device	\Device\HarddiskVolume4
Locale	United States
Hardware Abstraction Layer	Version = "10.0.22621.2506"
User Name	LAPTOP-DE1VAHVT\ABIKJITH REDDY
Time Zone	India Standard Time
Installed Physical Memory (RA	8.00 GB
Total Physical Memory	7.75 GB
Available Physical Memory	473 MB
Total Virtual Memory	23.2 GB

Figure 1: System confugration

## (i) Device specifications

Device name	LAPTOP-DE1VAHVT
Processor	11th Gen Intel(R) Core(TM) i5-11300H @ 3.10GHz 3.11 GHz
Installed RAM	8.00 GB (7.75 GB usable)
Device ID	9E65838C-7679-4C0C-AF1E-2B25F2E31975
Product ID	00327-36299-48467-AAOEM
System type	64-bit operating system, x64-based processor
Pen and touch	No pen or touch input is available for this display

#### Figure 2: hardware specification

Windows specific	ations
Edition	Windows 11 Home Single Language
Version	23H2
Installed on	02-11-2023
OS build	22631.2715
Experience	Windows Feature Experience Pack 1000.22677.1000.0

Figure 3: windows specification

## 3 Software and language required to Run this code

Software: Colab Pro Language: Python

## 4 Python Libraries required to Run this code

Figure 4, Figure 5, Figure 6 almost have all the libraries required for this project but one might need to install some of those libraries using pip install command before using and also might need libraries for any final estimators in stacking ensemble model.







Figure 5: libraries required for RoBERT and Staking ensemble model of BERT and RoBERT

### 5 Dataset used in this project

This project uses only 1 dataset which is Telugu news dataset which has 5 columns of which only 2 columns will be used for this project namely body and topic. Body has the news information while topic has labels to which that news belons to and there are 5 unique labels.

## 6 Code implementation snippet

- 6.1 BERT
- 6.2 RoBERT

#### 6.3 LSTM

#### 6.4 Stacking ensemble

Figure 14 shows major part of the code in order run these codes required libraries must be installed which are specified in report in the libraries section.

All the implementation codes above are the major parts of the code but in order to run those one might have to do some basic coding like splitting data and data loaders



Figure 6: libraries required for LSTM

	body	topic
0	భారీ ఎత్తున మొండిబకాయిలు పెరిగిపోవడంతో ఐడిబిఐ	business
1	న్యూఢిల్లీ : ఆర్థిక మంత్రి అరుణ్ జైట్లీ సోమవా	business
2	కటక్: ఇంగ్లండ్తో జరుగుతున్న సెకండ్ వన్డే మ్యా	sports
3	\nఇస్లామాబాద్ : పాకిస్థాన్ అంతర్జాతీయ ఉగ్రవాది	nation
4	స్టార్ హీరోగా వరుస సినిమాలతో బిజీగా ఉన్న ప్పటి	entertainment

Figure 7: Telugu news dataset

and also install required libraries which are mentioned in earlier section of this report. Stacking ensemble code provided same for all the stacking model but one has to specifty the meta estimators and final estimators as below.

#### 6.5 Estimators for stacking ensemble model

In Figure 14, there are 3 model BERT, RoBERT and LSTM each of them have different meta classifiers. For BERT and RoBERT it is logistic regression and for LSTM it is random forest classifier and as an final estimator we have used logistic regression. It can be adopted to any other stacking model one such example is BERT+LSTM model which is shown below in figure.

Figure 16, shows the meta estimator for BERT is logistic regression, and for LSTM is random forest and the final estimator is again logistic regression. And these estimators can be changed.



Figure 8: Part 1 of BERT code



Figure 9: Part 2 of BERT code



Figure 10: Part 1 of RoBERT code



Figure 11: Part 2 of RoBERT code







Figure 13: LSTM code



Figure 14: Stacking ensemble model code



Figure 15: Final estimator code for BERT+RoBERT+LSTM model



Figure 16: Final estimator code for BERT+LSTM model