

High Level Design (HLD)

News Article Classifier

Revision No: 1.0

Last Date of Revision: 18/11/2023.

Document Version Control-

<u>Date Issued</u>	<u>Version</u>	<u>Description</u>	<u>Author</u>
18/11/2023	1.0	HLD – V- 1.0	Mahesh. A

Contents -

Description	Page No.
Document Version Control	2
1. Introduction	4
1.1 Necessity of HLD	4
2. General Description	4
2.1 Product Perspective	4
2.2 Problem Statement	4
2.3 Data Requirements	4
2.4 The format would as the below mentioned sample	6
2.5 Tools Required	7
3. Design Details	8
3.1 Process Flow-	8
3.2 Model Training and Evaluation-	8
3.3 Deployment Process	9
3.4 Event Log	9
3.5 Error Handling	9
4 Performance-	
4.1 Reusability	10
4.2 Application Compatibility	10
4.3 Resource Utilization	10
4.4 Deployment	10
5 Conclusion	

1 Introduction

1.1 Necessity of HLD-

The purpose of this document is to add the necessary information to the current project for Modelling and its coding. This can be treated as a user manual.

2. General Description

2.1 Product Perspective-

This news article classifier is the project is based on the Machine learning Classifier which helps us to identify the news article class.

2.2 Problem S-

To create the AI solution for the news article classification

. To detect to news article type

2.3 Data Requirements-

The data requirements completely depend on the problem statement.

We need the data in the .CSV or any other pandas readable format with the below mentioned details.

Attribute Information: (classes: ['business', 'entertainment', 'politics', 'sport', 'technology'])

2.4 The format would as the below mentioned sample

1833	<p>worldcom ex-boss launches defence lawyers defending former WorldCom chief bernie ebbers against a battery of fraud charges have called a company whistleblower as their first witness. cynthia cooper worldcom s ex-head of internal accounting alerted directors to irregular accounting practices at the us telecoms giant in 2002. her warnings led to the collapse of the firm following the discovery of an \$11bn (Â£5.7bn) accounting fraud. mr ebbers has pleaded not guilty to charges of fraud and conspiracy. prosecution lawyers have argued that mr ebbers orchestrated a series of accounting tricks at worldcom ordering employees to hide expenses and inflate revenues to meet wall street earnings estimates. but ms cooper who now runs her own consulting business told a jury in new york on wednesday that external auditors arthur andersen had approved worldcom s accounting in early 2001 and 2002. she said andersen had given a green light to the procedures and practices used by worldcom. mr ebber s lawyers have said he was unaware of the fraud arguing that auditors did not alert him to any problems. ms cooper also said that during shareholder meetings mr ebbers often passed over technical questions to the company s finance chief giving only brief answers himself. the prosecution s star witness former worldcom financial chief scott sullivan has said that mr ebbers ordered accounting adjustments at the firm telling him to hit our books . however ms cooper said mr sullivan had not mentioned anything uncomfortable about worldcom s accounting during a 2001 audit committee meeting. mr ebbers could face a jail sentence of 85 years if convicted of all the charges he is facing. worldcom emerged from bankruptcy protection in 2004 and is now known as mci. last week mci agreed to a buyout by verizon communications in a deal valued at \$6.75bn.</p>	business
154	<p>german business confidence slides german business confidence fell in february knocking hopes of a speedy recovery in europe s largest economy. munich-based research institute ifo said that its confidence index fell to 95.5 in february from 97.5 in january its first decline in three months. the study found that the outlook in both the manufacturing and retail sectors had worsened. observers had been hoping that a more confident business sector would signal that economic activity was picking up. we re surprised that the ifo index has taken such a knock said dz bank economist bernd weidensteiner. the main reason is probably that the domestic economy is still weak particularly in the retail trade. economy and labour minister wolfgang clement called the dip in february s ifo confidence figure a very mild decline . he said that despite the retreat the index remained at a relatively high level and that he expected a modest economic upswing to continue. germany s economy grew 1.6% last year after shrinking in 2003. however the economy contracted by 0.2% during the last three months of 2004 mainly due to the reluctance of consumers to spend. latest indications are that growth is still proving elusive and ifo president hans-werner sinn said any improvement in german domestic demand was sluggish. exports had kept things going during the first half of 2004 but demand for exports was then hit as the value of the euro hit record levels making german products less competitive overseas. on top of that the unemployment rate has been stuck at close to 10% and manufacturing firms including daimlerchrysler siemens and volkswagen have been negotiating with unions over cost cutting measures. analysts said that the ifo figures and germany s continuing problems may delay an interest rate rise by the european central bank. eurozone interest rates are at 2% but comments from senior officials have recently focused on the threat of inflation prompting fears that interest rates may rise.</p>	business
1101	<p>bbc poll indicates economic gloom citizens in a majority of nations surveyed in a bbc world service poll believe the world economy is worsening. most respondents also said their national economy was getting worse. but when asked about their own family s financial outlook a majority in 14 countries said they were positive about the future. almost 23 000 people in 22 countries were questioned for the poll which was mostly conducted before the asian tsunami disaster. the poll found that a majority or plurality of people in 13 countries</p>	business

	<p>believed the economy was going downhill compared with respondents in nine countries who believed it was improving. those surveyed in three countries were split. in percentage terms an average of 44% of respondents in each country said the world economy was getting worse compared to 34% who said it was improving. similarly 48% were pessimistic about their national economy while 41% were optimistic. and 47% saw their family's economic conditions improving as against 36% who said they were getting worse. the poll of 22 953 people was conducted by the international polling firm globescan together with the program on international policy attitudes (pipa) at the university of maryland. while the world economy has picked up from difficult times just a few years ago people seem to not have fully absorbed this development though they are personally experiencing its effects said pipa director steven kull. people around the world are saying: i'm ok but the world isn't. there may be a perception that war terrorism and religious and political divisions are making the world a worse place even though that has not so far been reflected in global economic performance says the bbc's elizabeth blunt. the countries where people were most optimistic both for the world and for their own families were two fast-growing developing economies china and india followed by indonesia. china has seen two decades of blistering economic growth which has led to wealth creation on a huge scale says the bbc's louisa lim in beijing. but the results also may reflect the untrammelled confidence of people who are subject to endless government propaganda about their country's rosy economic future our correspondent says. south korea was the most pessimistic while respondents in italy and mexico were also quite gloomy. the bbc's david willey in rome says one reason for that result is the changeover from the lira to the euro in 2001 which is widely viewed as the biggest reason why their wages and salaries are worth less than they used to be. the philippines was among the most upbeat countries on prospects for respondents' families but one of the most pessimistic about the world economy. pipa conducted the poll from 15 november 2004 to 3 january 2005 across 22 countries in face-to-face or telephone interviews. the interviews took place between 15 november 2004 and 5 january 2005. the margin of error is between 2.5 and 4 points depending on the country. in eight of the countries the sample was limited to major metropolitan areas.</p>	
1976	<p>lifestyle governs mobile choice faster better or funkier hardware alone is not going to help phone firms sell more handsets research suggests. instead phone firms keen to get more out of their customers should not just be pushing the technology for its own sake. consumers are far more interested in how handsets fit in with their lifestyle than they are in screen size onboard memory or the chip inside shows an in-depth study by handset maker ericsson. historically in the industry there has been too much focus on using technology said dr michael bjorn senior advisor on mobile media at ericsson's consumer and enterprise lab. we have to stop saying that these technologies will change their lives he said. we should try to speak to consumers in their own language and help them see how it fits in with what they are doing he told the bbc news website. for the study ericsson interviewed 14 000 mobile phone owners on the ways they use their phone. people's habits remain the same said dr bjorn. they just move the activity into the mobile phone as it's a much more convenient way to do it. one good example of this was diary-writing among younger people he said. while diaries have always been popular a mobile phone -- especially one equipped with a camera -- helps them keep it in a different form. youngsters' use of text messages also reflects their desire to chat and keep in contact with friends and again just lets them do it in a slightly changed way. dr bjorn said that although consumers do what they always did but use a phone to do it the sheer variety of what the new handset technologies make possible does gradually drive new habits and lifestyles. ericsson's research has shown that consumers divide into different tribes that use phones in different ways. dr bjorn said groups dubbed pioneers and materialists were most interested in trying new things and were behind the start of many trends in phone use. for instance he said older people are using sms much more than they did five years ago. this was because younger users often the children of ageing mobile owners encouraged older people to try it so they could keep in touch. another factor</p>	tech

2.5 Tools Required-

Python Programming language and the frameworks below mentioned are used to build the model.

- Pandas
- Numpy
- Matplotlib
- Sklearn
- Logging
- Flask
- OS
- String
- NLTK
- SpaCy
- Tensorflow
- Keras
- Pickle
- Joblib

- a. VS code is used as the IDE
- b. For visualization we use Mat plot library or sea born
- c. We use local host for deploying the model
- d. HTML/CSS for the front end development
- e. Python flask used of the backend development
- f. Git is used for version control
- g. Docker is used of creation of model that works on any architecture.



XGBoost



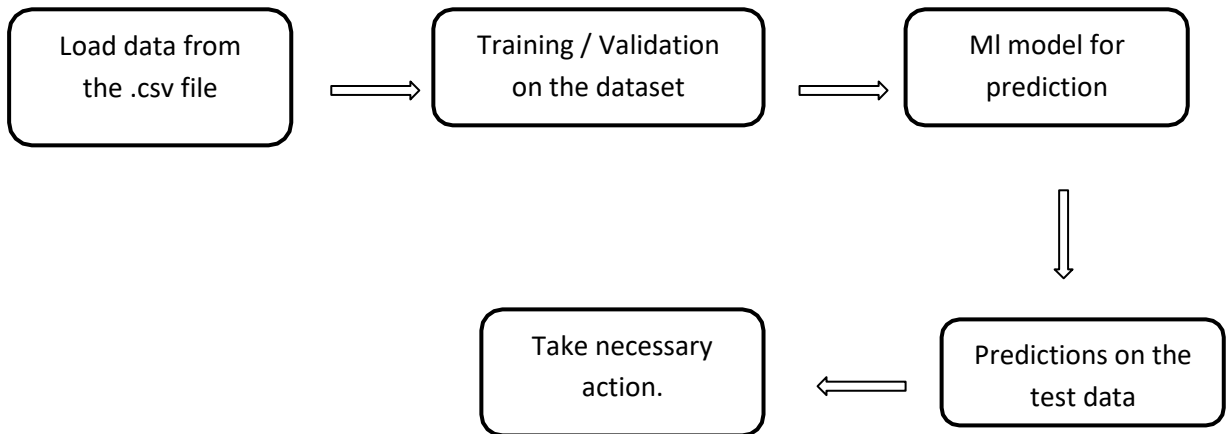
matplotlib

spaCy

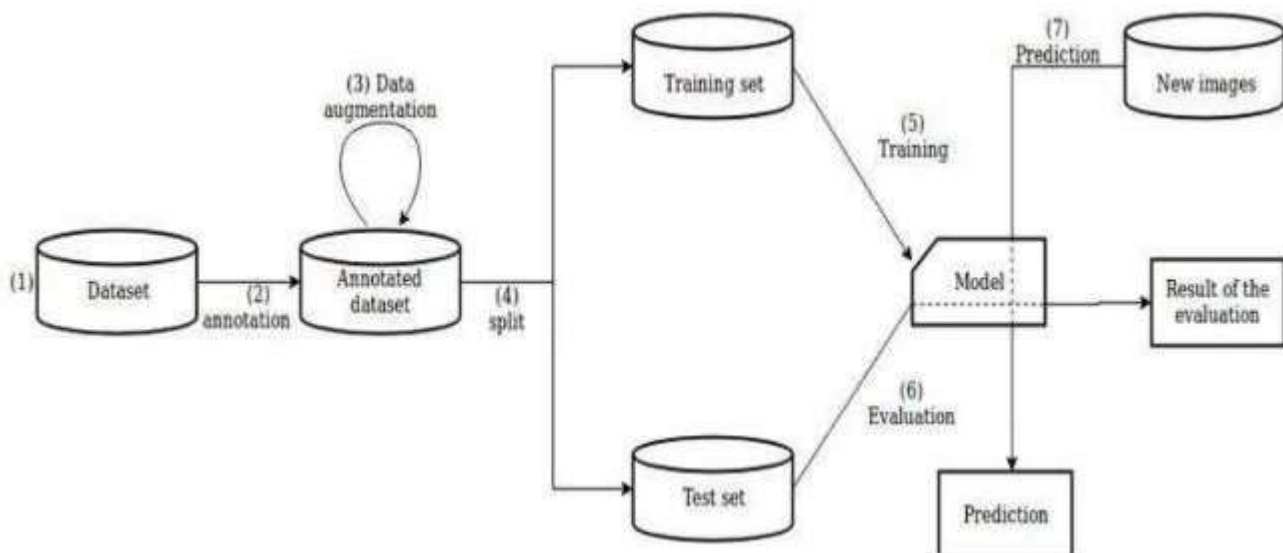


3 Design Details

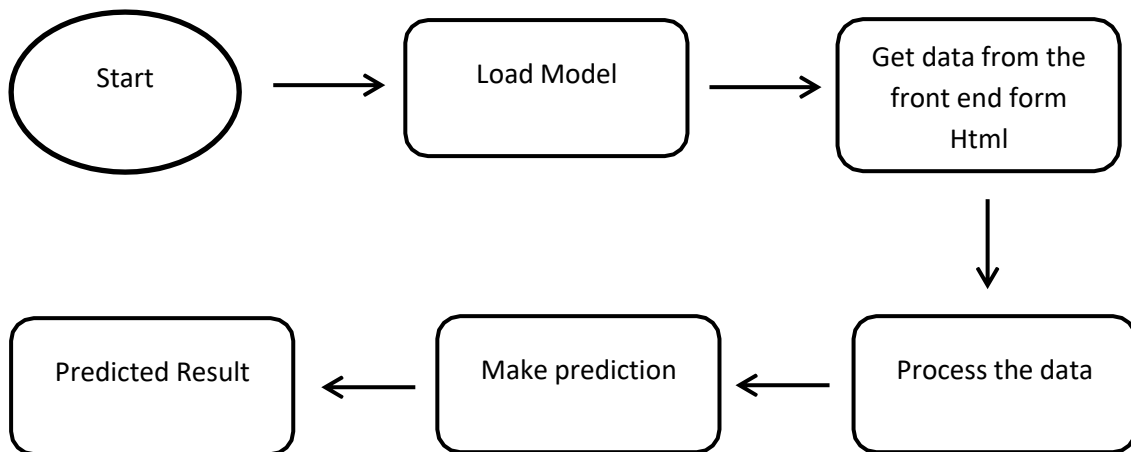
3.1 Process Flow-



3.2 Model Training and Evaluation-



3.3 Deployment Process:



3.4 Event Log-

The system should log every event so the user will know the process that happens on the time.

The system should identify the different logs

The system should note down the logs for further usage

Developer uses this data

3.5 Error Handling-

By using the logs file all errors can be noted and find the solution for them by the developer.

4 Performance-

This News article classifier ML model is used to generate the output to the class it belongs to.

4.1 Reusability-

The code written should be reused without any problem.

4.2 Application Compatibility-

We are using python as an interface.

4.3 Resource Utilization-

When any task is performed, it will use all the processing power to do the task assigned.

4.4 Deployment-



5 Conclusion-

This ML news article classifier which helps user to identify the category of the article it belongs to (['business', 'entertainment', 'politics', 'sport', 'technology'])