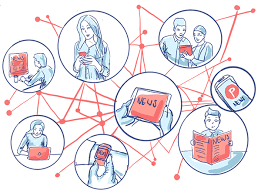
**Low Level Design (LLD)**

News Articles Sorting



**Revision Number - 1.0**

**Last Date of Revision – 07/07/2022**

**SRIPHANI**

# Document Control

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Date** |  | **Version** | **Description** | **Author** |
| 07/07/2022 | 1.0 |  | LLD document | Sriphani |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Contents**

Document Version Control…………………………………………………………………………2

1. Introduction……………………………………………………………………………………….3
   1. What is Low Level Design Document?.................................................................3
   2. Scope……………………………………………………………………………………4
   3. Project Introduction……………………………………………………………………4
2. Problem Statement.……………………………………………………………………………...5
3. Dataset Information……………………………………………………………………….……...5
4. Architecture…………………………………………………………………………………….…6
   1. Architecture Description …………………………………………….…....................6

# 1. Introduction

## 1.1 What is Low Level Design Document?

The goal of the Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Heart Disease Diagnostic Analysis dashboard. LLDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

**1.2 What is Scope?**

Low-level design (LLD) is a component-level design process that follows a stepby-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

## 1.3 Project Introduction

* A news article discusses current or recent news of either general interest (i.e. daily newspapers) or on a specific topic (i.e. political or trade news magazines, club newsletters, or technology news websites).
* A news article can include accounts of eyewitnesses to the happening event. We must have seen the news divided into categories when we go to a news website.
* Some of the popular categories that you’ll see on almost any news website are tech, entertainment, sports, etc. If you want to know how to classify news categories using machine learning, this article is for you.
* Every news website classifies the news article before publishing it so that every time visitors visit their website can easily click on the type of news that interests them. For example, I like to read the latest technology updates, so every time I visit a news website, I click on the technology section. But you may or may not like to read about technology, you may be interested in politics, business, entertainment, or maybe sports. Currently, the news articles are classified by hand by the content managers of news websites. But to save time, they can also implement a machine learning model on their websites that read the news headline or the content of the news and classifies the category of the news.

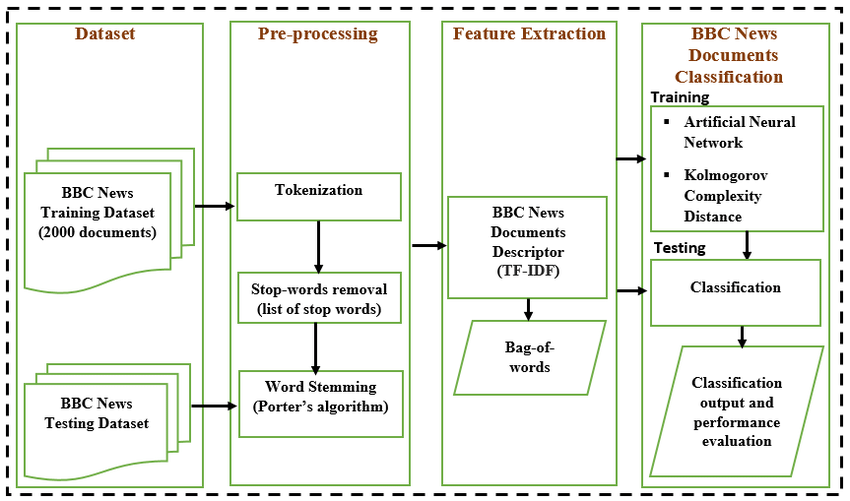
# 2. Problem Statement

* In today’s world, data is power.
* With News companies having terabytes of data stored in servers, everyone is in the quest to discover insights that add value to the organization.
* With various examples to quote in which analytics is being used to drive actions, one that stands out is news article classification.
* Nowadays on the Internet there are a lot of sources that generate immense amounts of daily news.
* In addition, the demand for information by users has been growing continuously, so it is crucial that the news is classified to allow users to access the information of interest quickly and effectively.
* This way, the machine learning model for automated news classification could be used to identify topics of untracked news and/or make individual suggestions based on the user’s prior interests

# 3. Dataset Information

* **Article Id** – Article id unique given to the record
* **Article** – Text of the header and article
* **Category** – Category of the article (tech, business, sport, entertainment, politics)

# 4. Architecture



## 4.1 Architecture Description

1. **Raw Data Collection** The Dataset was taken from iNeuron’s Provided Project Description Document.

[BBC News Classification | Kaggle](https://www.kaggle.com/c/learn-ai-bbc/data)

### 2. Data Pre-Processing

Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data feeded to the model to train.

This Process includes-

1. Handling Null/Missing Values
2. Handling Skewed Data
3. Outliers Detection and Removal

### 3. Data Cleaning

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

1. Remove duplicate or irrelevant observations
2. Filter unwanted outliers
3. Renaming required attributes

### 4. Exploratory Data Analysis (EDA)

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

### 5. Reporting

Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self-explanatory report because your model will be used by many stakeholders who are not from technical background.

1. High Level Design Document (HLD)
2. Low Level Design Document (LLD)
3. Architecture
4. Wireframe
5. Detailed Project Report
6. Power Point Presentation

### 6. Modelling

Data Modelling is the process of analysing the data objects and their relationship to the other objects. It is used to analyse the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

### 7. Deployment

We created a UI

