Project Name - Dont Use short names of project

Six months Training Report

SUBMITTED IN PARTIAL FULFILLMENT REQUIREMENT FOR

THE AWARD OF DEGREE OF

Bachelor of Technology

(Computer Science & Engineering)

SUBMITTED BY

NAME

(UNIVERSITY ROLL No. 1232434)

December, 2021

SABUDH

D-228, Third Floor, Sector 74,

Sahibzada Ajit Singh Nagar,

Punjab 140307

CANDIDATE'S DECLARATION

I hereby certify that I have under gone six months industrial training at SABUDH FOUNDATION and worked on project entitled, "**Project Name - Dont Use short names of project**", in partial fulfillment of requirements for the award of Degree of **Bachelor of Technology** in name of the department at **institute name**, having University Roll No.1232434, is an authentic record of my own work carried out during a period from July, 2021 to December, 2021 under the supervision of MENTOR NAME.

(**name)**

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(**MENTOR NAMES)**

Person from Organisation with designation

ABSTRACT

Type your abstract here

ACKNOWLEDGEMENT

I am highly grateful to **MENTOR NAME, DESIGNATION**, Sabudh Foundation, Mohali, for providing opportunity to carry out six months training at Sabudh Founadtion from July-December 2021.

(Name of the Mentor) has provided great help in carrying out my work and is acknowledged with reverential thanks. Without the wise counsel and able guidance, it would have been impossible to complete the training in this manner.

I would like to express thanks profusely to thank **MENTOR NAME**, for stimulating me time to time. I would also like to thank entire team of Sabudh Foundation. I would also thank my friends who devoted their valuable time and helped me in all possible ways towards successful completion of this training.

(name)

LIST OF FIGURES

Figure No.	Figure Description	Page No.

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
SVM	Support Vector Machine
abbreviation	Full Form
abbreviation	Full Form

TABLE OF CONTENTS

Chapter 1

Introduction to Organisation

Sabudh Foundation - An NGO that applies data science for social good. Sabudh Foundation is formed by the leading data scientists in the industry with the objective to bring together data and young data scientists to work on focused, collaborative projects for social benefit. Sabudh foundation is working on solving the real and high impact problems in areas such as education, governance, healthcare, and agriculture using Artificial Intelligence and Machine Learning techniques.

Data science can be used across a number of industries in order to be beneficial for the society. For example in agriculture, there are now Agrobots and drones being used to gauge the health of the harvest that can help farmers improve their crop yield and reduce costs. With the help of advanced technologies, we're able to save 90
The foundation has taken steps to involve Colleges, Universities, and Industry from the region for the social cause. Particularly, the foundation has signed academic and research-based MoUs with Panjab University, Chandigarh, GNDEC, Ludhiana, BML Munjal University, Punjab Government (Punjab Police), Punjabi University, Patiala, and Punjab Engineering College, Chandigarh.

Chapter 2

Introduction to Project

2.1 Overview

Software bug reports are an important source of information for software development and maintenance. During resource allocation, a project manager distributes time and effort for each project task and assigns appropriate developers.

It has been investigated that research scholars spend a significant time of total research period on identifying the current and promising research areas. Even after choosing a specific research area, the prevailing research trends are not always well understood. Thus, the research area selection is critical and a time-consuming process. The researchers perform manual literature reviews to reveal the research gaps, but they are often subjective and biased.

The core idea behind this work is to use an unsupervised technique (topic modeling) to identify the latent semantic patterns from the bibliographic database. It would answer the research questions viz. the prevailing trends and the topics which are required for research problem framing. The model automatically clusters the text documents into one or more defined number of categories. The model has been developed by using Python because of it's strong string class with powerful methods.

2.2 Existing System

2.3 User Requirement Analysis

2.4 Feasibility Study

Chapter 3

Literature Review

Add the summaries of research work you referred for your work. Either discuss the research in chronological order or by topic. Do cite all the research articles at the end of report in Bibliography section. For example the below paragraphs discuss the research in a flow of the topic

It has been very crucial task for researchers to identify the current and historical trends. Further, extraction from this information to focus on innovative methods and concepts also requires critical thinking. Keeping in view the volume of unstructured data, it is a tedious task to extract the desired information. But with technological advances, various methods have been developed for extracting the relevant information from voluminous data. The data from various bibliographic sources viz. research articles, book chapters, patents, and technical reports, can be summarized by identification of topics [1, 2].

Topic modeling is an automated process which is based on algorithmic-based analysis [3, 4]. From a corpus, this method is used to identify the patterns. Further, semantic meaning is also added to the corpus's vocabulary. Topic modeling can be applied by two methods viz. topic analysis and clustering but topic analysis is considered better choice for research trends' identification [5]. The main difference between the two methods is the assignment of a document to the topic or cluster. A document can be assigned to multiple topics in topic analysis, whereas it joins exactly one cluster in clustering.

3.1 Comparison

This section should have table where you need to compare the results from different prominent research of area of interest.

3.2 Objectives of Project (Must be clearly, precisely defined and Implementation must be done.)

This section should elaborate the objectives based on the gaps from the previous researchers.

Chapter 4

Exploratory Data Analysis

4.1 Dataset

Discussion on the dataset, scraping methods, pipelines etc.

4.2 Exploratory Data Analysis and Visualisations

Details analysis of the data and you could include feature engineering into this section

4.3 Related Sections

Chapter 5

Methodology

This section discusses the methodology to develop the ML model. It may include the following but not limited to:

5.1 Introduction to Languages (Front End and Back End)

5.2 Any other Supporting Languages/ packages

5.3 User characteristics

5.4 Constraints

5.5 Use Case Model/Flow Chart/DFDS

5.6 Database design

5.7 Table Structure

5.8 ER Diagrams

5.9 Assumptions and Dependencies

5.10 ML algorithm discussion

5.11 Implementation of Algorithm with Screen Shots/ Figures (Each Figure must be numbered and Description of Figure must be provided)

Chapter 6

Results

Chapter 7

Conclusion and Future Scope

7.1 Conclusion

7.2 Future Scope

Bibliography

[1] Q. Mei and C. Zhai, "Discovering evolutionary theme patterns from text- an exploration of temporal text mining," in *The Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Chicago, USA, Aug. 2005.

[2] J. L. Hurtado, A. Agarwal, and X. Zhu, "Topic discovery and future trend forecasting for texts," *Journal of Big Data*, vol. 3, no. 7, pp. 1–21, 2016.

[3] K. R. Canini, L. Shi, and T. L. Griffiths, "Online inference of topics with latent dirichlet allocation." *Journal of Machine Learning Research*, vol. 5, pp. 65–72, 2009. [Online]. Available: http://jmlr.org/proceedings/papers/v5/canini09a/canini09a.pdf

[4] S. Saini, B. Kasliwal, and S. Bhatia, "Language identification using g-lda," *International Journal of Research in Engineering and Technology*, vol. 2, no. 11, pp. 42–45, nov 2013. [Online]. Available: https://issuu.com/ijret/docs/language_identification_using_g-lda

[5] N. Evangelopoulos, X. Zhang, and V. R. Prybutok, "Latent semantic analysis: five methodological recommendations," *European Journal of Information Systems*, vol. 21, no. 1, pp. 70–86, 2012.

Appendix (Any additional Information regarding Project)