

KGiSL Institute of Technology

NAAN MUDHALVAN

Project title :

Public Health Awareness

TEAM MEMBERS:

* Shree S M
* Hentry Abel
* Hariprasath
* Ajay Raahul P

PROJECT DESCRIPTION :

**Phase 2: INOVATION & DESIGN**

**1.**\*\*Define Clear Objectives\*\*:

* SWOT Analysis: Conduct a SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis
* Surveys and Questionnaires
* Data Visualization Tools
* Google Trends
* Social Media Analytics Tools
* Data Science Frameworks
* Project Management Tools

2. \*\*Data Collection\*\*:

* Kaagle ( dataset Provided )
* Web Analytics
* Mobile Apps and Health Tracking
* Community Surveys and Health Records
* Geographic Information System (GIS) Mapping
* Health Records and Databases

3. \*\*Data Cleaning and Preparation\*\*:

Technique:

* Data Profiling
* Handling Missing Values
* Duplicate Data
* Outlier Detection and Treatment
* Data Transformation
* Date and Time Handling
* Data Scaling
* Text Data Cleaning
* Data Integration:

4. \*\*Data Analysis\*\*:

TOOLS USED:

* Tableau or Power BI
* Python or R
* SAS(Statistical Analysis System)
* Open Refine
* Trifacta
* SQL
* Excel
* JUPYTER or Google Colab NoteBooks

5. \*\*Visualization\*\*:

* Tableau
* Power BI

6. \*\*python libraries \*\*:

* Pandas: For data manipulation and cleaning.
* NumPy: For numerical operations and handling missing values.
* Scikit-learn: For preprocessing, including feature scaling and encoding.
* NLTK or SpaCy: For text data cleaning and preprocessing.

7.\*\* Implementation\*\*:

* The implementation phase involves coding the machine learning model, conducting data analysis, and creating a user-friendly interface for users to input campaign details and receive predictions.
* The implementation can be carried out using popular data science libraries such as Python's scikit-learn and TensorFlow.

8.\*\* Evaluation\*\*:

* The success of our solution will be evaluated based on the model's predictive accuracy and its ability to assist organizations in making informed decisions about future mental health awareness campaigns.
* The model's performance will be monitored and fine-tuned as new campaign data becomes available.

9.\*\* Proposed Solution\*\*:

We propose building a machine learning model that uses historical data to predict the success of future mental health awareness campaigns. The success of a campaign will be defined based on key performance indicators (KPIs) such as:

* Participation rate:

The percentage of employees who actively engage with the campaign.

* Satisfaction level:

Feedback from employees regarding the campaign's impact on their mental health.

* Retention rate:

The percentage of employees who remain with the organization after the campaign.

10. Conclusion:

* By incorporating machine learning algorithms to predict the success of mental health awareness campaigns, organizations in the tech industry can make data-driven decisions, allocate resources more effectively, and ultimately improve the well-being of their employees. This solution will contribute to a more supportive and mentally healthy workplace.

11. References

* Kaggle Dataset: Mental Health in Tech Survey. Retrieved from <https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey>

This design document outlines our approach to using machine learning for predicting the success of mental health awareness campaigns in the tech industry. The implementation of this solution will require collaboration with data scientists, access to historical data, and a commitment to data-driven decision-making.