MindWell: Predictive Wellness for Professionals

Milestone 1: Project Initialization and Planning Phase

The "Project Initiation and Planning Phase" marks the beginning of the mental health prediction project for working professionals, where goals, scope, and stakeholders are defined. This critical phase establishes project parameters, identifies essential team members, allocates resources, and outlines a realistic timeline. It also involves assessing risks and planning strategies to mitigate them. Successful initiation sets the foundation for a well-organized and efficiently executed project, ensuring clarity, alignment, and proactive measures for potential challenges in predicting mental health conditions among working professionals.

Activity 1: Define Problem Statement
Problem Statement: Working professionals need a reliable way to assess their mental health status and determine if they should seek treatment or support. They require a system that can accurately predict whether they are likely to benefit from mental health resources based on their personal information and circumstances.

Ref. template: Click Here
MindWell Problem Statement Report: Click Here

Activity 2: Project Proposal (Proposed Solution)
This project aims to leverage Machine Learning to create a screening tool for mental health concerns in the workplace. The system will gather user input through surveys or questionnaires and analyze it using various algorithms (Logistic Regression, KNN, Decision Tree, etc.). Based on the analysis, the system will generate a prediction about the user's potential need for mental health services.
The project will include:
- Data collection and randomization of relevant mental health surveys from working professionals.
- Training and evaluation of Machine Learning models to identify themost accurate predictor.
- Development of a user interface for interaction with the system.
- Deployment of the model in a secure and user-friendly manner.

Ref. template: Click Here
SmartLender Project Proposal Report: Click Here

Activity 3: Initial Project Planning
During the initial project planning phase of the mental health prediction project for working professionals, key objectives are outlined, scope is defined, and stakeholders are identified. This phase involves setting timelines, allocating resources, and determining the overall project strategy. The team establishes a clear understanding of the dataset, formulates analysis goals, and plans the workflow for data processing. Effective initial planning sets the foundation for a systematic and well-executed project, ensuring successful outcomes. Additionally, the project follows an agile methodology with multiple sprints. Each sprint consists of functional requirements (epics), user stories with assigned story points and priorities, and team members responsible for completing tasks. Sprint start and end dates are also planned to ensure timely delivery of project milestones.

Ref. template: [Click Here](#)
SmartLender Project Planning Report: [Click Here](#)

Milestone 2: Data Collection and Preprocessing Phase

In the initial phase of Data Collection and Preprocessing for the mental health prediction project among working professionals, a plan is executed to collect pertinent data from appropriate sources, ensuring its quality through verification and handling missing values. Preprocessing activities encompass cleaning, encoding, and structuring the dataset to facilitate subsequent exploratory analysis and development of machine learning models.

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

The 'Mental Health in Tech Survey' dataset found on Kaggle, this dataset contains information pertinent to predicting mental health conditions among working professionals. It encompasses various attributes relevant to mental health assessment and undergoes rigorous data preprocessing, including handling missing values and ensuring ethical compliance. This meticulous process guarantees a dependable groundwork for developing accurate predictive models for mental health prediction in the workplace.

Ref. template: [Click Here](#)
SmartLender Data Collection Report: [Click Here](#)

Activity 2: Data Quality Report

The Data Quality Report Template outlines issues and solutions for a dataset. It identifies high-severity problems like incorrect age values, proposing removal of erroneous entries. Moderate-severity concerns, such as uneven country distribution, suggest column removal or alternative grouping methods. Missing data points in certain columns prompt moderate-severity imputation using the mode. Low-severity issues, like inconsistent gender formatting, suggest standardization into common categories for improved consistency. Consideration for an "Other" category is advised for significant variations beyond the standard options.

Ref. template: [Click Here](#)
SmartLender Data Quality Report: [Click Here](#)

Activity 3: Data Exploration and Preprocessing

In the context of building a mental health prediction system, the template delineates steps for data exploration and preprocessing. It underscores the significance of resolving data quality issues, exploring data distributions, performing feature engineering, and splitting data for model development. The dataset, likely structured as a CSV file, encompasses survey responses covering demographics, work environment, and mental health experiences. Univariate analysis entails grasping variable types and descriptive statistics, while bivariate and multivariate analyses necessitate access to the actual data for probing relationships between variables. The section delves into techniques for identifying and addressing outliers, underscoring prudence and domain knowledge, particularly concerning sensitive mental health data.

Ref. template: [Click Here](#)
SmartLender Data Exploration and Preprocessing Report: [Click Here](#)

Milestone 3: Model Development Phase

The Model Development Phase involves creating a predictive model for mental health prediction among working professionals. It includes strategic feature selection, evaluating and choosing models such as Random Forest, Decision Tree, KNN, and XGB. The phase initiates model training with code implementation and extensively validates and assesses model performance to make informed decisions regarding mental health predictions in the workplace.

Activity 1: Feature Selection Report

In the upcoming update, each feature will come with a brief description. Users will select whether it's included, providing reasoning for their decision, streamlining decision-making, and enhancing transparency in feature selection. Key features include Age, Gender, and treatment-related questions, while redundant or potentially biased features like Timestamp and Country will be excluded. This process ensures relevant data inclusion and minimizes bias in the mental health prediction model for working professionals.

Ref. template: [Click Here](#)
SmartLender Feature Selection Report: [Click Here](#)

Activity 2: Model Selection Report

The Model Selection Report rationalizes the selection of Random Forest, Decision Tree, KNN, and XGB models for loan approval prediction. It assesses their strengths in managing intricate relationships, interpretability, adaptability, and predictive accuracy, ensuring alignment with project goals.

Ref. template: [Click Here](#)
SmartLender Model Selection Report: [Click Here](#)

Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

In the forthcoming demonstration, the initial model training code will be displayed via a screenshot. Additionally, the model validation and evaluation report will encompass classification reports, accuracy metrics, and confusion matrices for various models, each presented through corresponding screenshots..

Ref. template: [Click Here](#)
SmartLender Model Development Phase Template: [Click Here](#)

Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Activity 1: Hyperparameter Tuning Documentation

The XGB Classifier model stood out due to its exceptional performance, showcasing impressive accuracy throughout the hyperparameter tuning process. Its capability to manage intricate relationships, mitigate overfitting, and enhance predictive accuracy are in line with project goals, validating its choice as the ultimate model.

Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Report delineates the disparities between baseline and optimized metrics across different models, emphasizing the notable improvement in performance observed with the XGB Classifier model. This analysis elucidates the refined predictive prowess attained through meticulous hyperparameter tuning, offering valuable insights into model enhancement.

Activity 3: Final Model Selection Justification

The Final Model Selection Justification articulates the rationale for choosing XGB Classifier as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal loan approval predictions.

Ref. Template: Click Here
SmartLender Model Optimization and Tuning Phase Report: Click Here

Milestone 5: Project Files Submission and Documentation

For project file submission in Github, Kindly click the link and refer to the flow. **Click Here**
For the documentation, Kindly refer to the link. **Click Here**

Milestone 6: Project Demonstration

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.