LLM Experimentation: Enhancing Mental Health Predictions with Natural Language Explanations

Objective:

The goal of this experimentation is to utilize a free or open-source Large Language Model (LLM) to provide human-friendly explanations for mental health condition predictions. Additionally, the LLM is used to suggest coping mechanisms and potential next steps for individuals based on their responses.

Implementation:

1. Model Selection

For this experiment, we integrated **Google's Gemini 1.5 Pro**, a state-of-the-art generative AI model, to generate natural language explanations based on the prediction results. The model was chosen due to its ability to understand contextual information and generate detailed, meaningful responses.

2. Workflow

1. **User Input Collection**: The user provides responses to various features related to their mental health.
2. **Preprocessing**:

   - Binary categorical variables (e.g., Yes/No) are mapped to numerical values (0/1).
   - Ordinal categorical variables are encoded based on predefined rankings.
   - The input data is standardized using a pre-trained scaler.

3. **Prediction using Random Forest**:

   - The selected features are fed into the trained **Random Forest model**.
   - The model outputs a binary prediction: **Will Seek Treatment** or **Will Not Seek Treatment**.
   - A confidence score (probability of prediction) is also provided.

4. **Reversing Encoded Features**:

   - To improve interpretability, the encoded features are mapped back to their original categorical values.

5. **Generating Natural Language Explanations**:

   - A structured **prompt** is created using the reversed feature values and the model's prediction.

Prompt given:

"""

A person with the following mental health assessment features has been analyzed:

{mapped_input_selected.to_dict(orient='records')[0]}

Based on this assessment, the model predicts that this individual {'will seek treatment' if prediction[0] == 1 else 'will not seek treatment'}.

Provide:

1. A detailed and descriptive natural language explanation for the prediction.

2. Suggested coping mechanisms and potential next steps.

"""

- The **Gemini 1.5 Pro model** generates a human-readable explanation and suggests coping mechanisms.

6. **Output Display**:

- The final results, including the prediction, confidence score, and LLM-generated insights, are presented to the user.

Example Output:

User Input :

```
Enter value for Age: 25
Choose one for self_employed: Yes/No
Enter value for self_employed: Yes
Choose one for family_history: Yes/No
Enter value for family_history: Yes
Choose one for work_interfere: Never, Rarely, Sometimes, Often
Enter value for work_interfere: Sometimes
Choose one for no_employees: 1-5, 6-25, 26-100, 100-500, 500-1000, More than 1000
Enter value for no_employees: 26-100
Choose one for remote_work: Yes/No
Enter value for remote_work: Yes
Choose one for tech_company: Yes/No
Enter value for tech_company: Yes
Choose one for benefits: No, Don't know, Yes
Enter value for benefits: Yes
Choose one for care_options: No, Not sure, Yes
Enter value for care_options: Yes
Choose one for wellness_program: No, Don't know, Yes
Enter value for wellness_program: Yes
Choose one for seek_help: No, Don't know, Yes
Enter value for seek_help: Yes
Choose one for anonymity: No, Don't know, Yes
Enter value for anonymity: Yes
Choose one for leave: Very difficult, Somewhat difficult, Don't know, Somewhat easy, Very easy
Enter value for leave: Somewhat easy
Choose one for mental_health_consequence: No, Maybe, Yes
Enter value for mental_health_consequence: Yes
Choose one for phys_health_consequence: No, Maybe, Yes
Enter value for phys_health_consequence: Yes
Choose one for coworkers: No, Some of them, Yes
```

```
Choose one for coworkers: No, Some of them, Yes
Enter value for coworkers: Yes
Choose one for supervisor: No, Some of them, Yes
Enter value for supervisor: Yes
Choose one for mental_health_interview: No, Maybe, Yes
Enter value for mental_health_interview: Yes
Choose one for phys_health_interview: No, Maybe, Yes
Enter value for phys_health_interview: Yes
Choose one for mental_vs_physical: No, Don't know, Yes
Enter value for mental_vs_physical: Yes
Choose one for obs_consequence: Yes/No
Enter value for obs_consequence: Yes


  ◆  **Mental Health Treatment Prediction**  ◆
Prediction: Will Seek Treatment
Confidence Score: 0.91
```

🟣 **Explanation from Gemini:**
1. Detailed Explanation for the Prediction:

The model predicts that this individual will seek treatment due to several factors identified in their mental health assessment:

* **Positive Symptoms:**
    * Family history of mental health issues
    * Work performance is sometimes affected
    * Experienced negative mental health consequences
* **Resources and Support:**
    * Care options are available
    * Taking leave from work is somewhat easy
    * Has access to health and wellness benefits
* **Social Support:**
    * Has coworkers and a supervisor who are supportive
* **Help-Seeking Behavior:**
    * Recognizes the need for professional help and expresses a willingness to seek it
    * Participates in a workplace wellness program

The combination of these factors suggests that the individual is experiencing significant mental health challenges that are impacting their well-being and functioning. They have recognized the need for help and have

2. Suggested Coping Mechanisms and Potential Next Steps:

* **Seek Professional Help**: Encourage the individual to schedule an appointment with a mental health professional, such as a therapist or psychiatrist.
* **Utilize Available Resources**: Make sure the individual is aware of the care options and benefits available to them, such as employee assistance programs or health insurance coverage.
* **Implement Self-Care Practices**: Recommend strategies for managing stress, such as exercise, meditation, or spending time in nature.
* **Consider Workplace Accommodations**: Discuss with the individual and their supervisor the possibility of workplace accommodations that may reduce stress or improve well-being, such as flexible work arrangements
* **Encourage Social Support**: Emphasize the importance of connecting with supportive friends, family members, or peers who can provide emotional support and understanding.
* **Follow Up and Monitoring**: Schedule regular follow-up appointments with the individual to monitor their progress, provide additional support as needed, and adjust coping mechanisms as necessary.

Model Prediction:

- **Prediction**: Will Seek Treatment
- **Confidence Score**: 0.91

LLM-Generated Explanation & Suggestions:

Explanation:

😔 **Explanation from Gemini:**
1. Detailed Explanation for the Prediction:

The model predicts that this individual will seek treatment due to several factors identified in their mental health assessment:

* **Positive Symptoms:**
    * Family history of mental health issues
    * Work performance is sometimes affected
    * Experienced negative mental health consequences
* **Resources and Support:**
    * Care options are available
    * Taking leave from work is somewhat easy
    * Has access to health and wellness benefits
* **Social Support:**
    * Has coworkers and a supervisor who are supportive
* **Help-Seeking Behavior:**
    * Recognizes the need for professional help and expresses a willingness to seek it
    * Participates in a workplace wellness program

The combination of these factors suggests that the individual is experiencing significant mental health challenges that are impacting their well-being and functioning. They have recognized the need for help and have access to appropriate resources.

Coping Mechanisms & Next Steps:

2. Suggested Coping Mechanisms and Potential Next Steps:

* **Seek Professional Help**: Encourage the individual to schedule an appointment with a mental health professional, such as a therapist or psychiatrist.
* **Utilize Available Resources**: Make sure the individual is aware of the care options and benefits available to them, such as employee assistance programs or health insurance coverage.
* **Implement Self-Care Practices**: Recommend strategies for managing stress, such as exercise, meditation, or spending time in nature.
* **Consider Workplace Accommodations**: Discuss with the individual and their supervisor the possibility of workplace accommodations that may reduce stress or improve well-being, such as flexible work arrangements or reduced workload.
* **Encourage Social Support**: Emphasize the importance of connecting with supportive friends, family members, or peers who can provide emotional support and understanding.
* **Follow Up and Monitoring**: Schedule regular follow-up appointments with the individual to monitor their progress, provide additional support as needed, and adjust coping mechanisms as necessary.

Findings & Insights:

- The LLM successfully translated model predictions into **meaningful explanations**, making it easier for non-technical users to understand why a particular prediction was made.
- The suggested coping mechanisms provide actionable insights, **bridging the gap between AI-driven predictions and real-world mental health support**.
- Using an LLM to generate human-readable explanations enhances **interpretability**, which is critical for applications in sensitive domains like mental health.

Conclusion:

The integration of a Large Language Model (LLM) with predictive models in mental health analysis offers **significant advantages** in terms of user interpretability and actionable insights. The ability to explain model predictions in **natural language** and suggest next steps **enhances trust and usability**, making AI-driven mental health assessments more accessible to a wider audience.