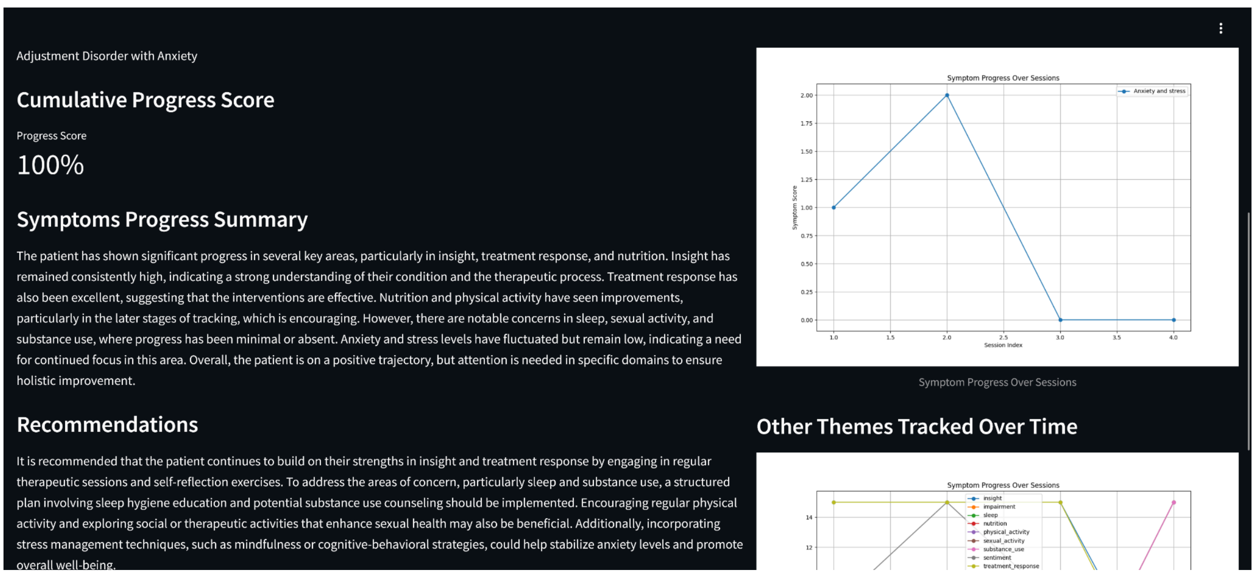
Documentation: Patient Progress Tracking System

Introduction

This system leverages AI, mathematical scoring mechanisms, and advanced data processing techniques to track a patient's progress over multiple therapy sessions. It is designed to provide actionable insights by scoring symptoms, standardizing symptom names across sessions, computing cumulative progress, and generating visualizations, summaries and recommendations. Below is a detailed documentation of the approach, including data flow, mathematical computations, and system components.



Sample image of solution.

Overview of Workflow

1. **File Processing**:
    - Therapy session notes (as provided in the email) in JSON or TXT format are uploaded.

- The files are parsed, and relevant fields are extracted to form a structured dataset.
- The fields extracted are all symptoms, insights, impairment and challenges, nutrition, sleep, substance use, quote (chief complaint), sexual activity, physical activity, response to treatment.

2. **Symptom Scoring**:
   - Symptoms and metrics are scored based on predefined scales inspired by PHQ–9.
   - Sentiment analysis is conducted to evaluate the tone of the patient's complaints.
3. **Symptom Standardization**:
   - Symptom names are standardized to eliminate redundancies/duplicates and ensure consistency across sessions.
4. **Tracking Metrics Over Time**:
   - Specific metric themes like insights, nutrition etc which can reveal a lot about patients progress and symptoms are tracked over multiple sessions, generating a time-series kind of data.
5. **Progress Scoring**:
   - Progress is calculated using weighted averages that pay more attention to more recent sessions, penalties for regressions, and a cumulative scoring system.
6. **Summarization and Recommendations**:
   - A summary of the patient's progress and tailored recommendations are generated using openai's chatgpt model. A temperature of 0 and seed of 42 are used for more consistent and deterministic results.
7. **Visualization**:
   - Time-series plots for symptoms and other metrics are generated and returned for easy visualization and interpretation.

All functions mentioned and discussed here can be found in the github repository as is. Therefore, it will be best if the repository is opened alongside this documentation for better and faster understanding.

File Processing

In this section, I make the assumption that therapy sessions have already been adequately documented into the SOAP format or session samples as shared in the email. Since an AI model had already transcribed and extracted these insights from the sessions earlier and also due to faster implementation time, it makes sense to continue working with this structured format for efficiency.

Uploaded files (json or text files) are processed using the `load_file` function, which parses and loads the content into dictionaries. Key fields such as symptoms, treatment response, impairments, and patient complaints are extracted recursively using the `extract_relevant_fields` function. This ensures all required fields are captured, even in nested structures.

The solution or implementation flow occurs in the exact fashion as described below.

Symptom Scoring

Approach

The system employs a large language model (LLM) to dynamically score symptoms and additional metrics such as treatment response, insight, and sleep. The scoring is based on:

- **Frequency and Severity for Symptoms**:
  - Daily: 0- 3
  - More than half the days: 4 - 7
  - Several days: 8 -11
  - None: 12 - 15

0 represents the worst symptom in terms of severity and frequency, 15 represents absence of that symptom (best state)

- **Extra Metrics** (Range: 0–15):

  Metrics such as insights, nutrition, response to treatment, sexual activity, substance use, sleep, physical activity are also measured and tracked. These factors are also very important. For example, insight is a very serious factor during the estimation of clinical depression, nutrition and physical activity can tell about the patient's attitude towards his health, social relationships respectively. I use the guideline below:

- 0–5: Negative or concerning observations.
- 6–10: Neutral or moderate observations.
- 11–15: Positive or improving conditions.

- **Sentiment/Mood Analysis** (Range: -1 to +1):

  We score the tone of the patient using the 'Quote (chief complaint)' field. This can give us information about how the patient feels about his improvement, relationships and challenges.

  - Sentiment of the patient's complaint is assessed to gauge emotional state.

We add all these guidelines to the prompts. We also prompt the LLM (gpt-4) in this case, that If any parameter is missing (`N/A`), it should be set to the **best score** for that category.

Symptom Standardization

To ensure uniform symptom tracking across sessions, symptom names are standardized using a dictionary that maps redundant or synonymous names to unique identifiers. For instance:

```
{
  "Anxiety and stress": "Anxiety",
  "Anxious and stressed": "Anxiety",
  "poor sleep": "Sleep Issues",
  "inadequate sleep": "Sleep Issues"
}
```

This process eliminates inconsistencies and allows for accurate tracking across sessions.

Tracking Metrics Over Time

The `track_metrics_over_time` function then creates a dictionary of time-series data for symptoms, themes (other metrics mentioned above) and sentiments. Each session's scores for all metrics are appended to their respective keys (symptoms/metrics).

Example:

```
{
  "Anxiety": [0, 10, 8],
  "Hopelessness": [5, 10],
```

    "Sleep Issues": [3, 7, 15, 11],
    "insight": [11, 12, 15]
}

Progress Scoring

Mathematical Approach

Now, we have to find a way to score the progress of the patient. We can simply consider the difference in the score of the patient between the first and most recent session for each symptom or metric. However, this method completely ignores any progress the patient has made in between each session. For instance, a patient may have complained about poor sleep during his first session and then experience good sleep in subsequent sessions except for the most recent session.
Also, since we expect patients to make consistent progress from session to session, then emphasis should be placed on more recent sessions.

To make this more holistic, I make use of a different approach that makes use of weighted averages that weights the recent sessions more. I make use of a linear weighting system and something similar to a softmax computation. Each part is explained briefly below:

1. Weighted Average

For each symptom or metric, recent sessions are weighted more heavily to reflect the expectation of consistent improvement (n = number of therapy sessions had so far):

$$\text{Weighted Average} = \frac{\sum_{t=1}^{n} \text{Score at Time Step } t \times \text{Weight}_t}{\sum_{t=1}^{n} \text{Weight}_t}$$

Where:

$$\text{Weight}_t = \frac{t}{n}, \quad t \in [1, n]$$

From the above equation, we see that the weights allocated to a session get bigger as therapy continues.

2. Progress Percentage

Then for each symptom, progress is calculated relative to the initial score at first session:

$$\text{Progress Percentage} = \frac{\text{Weighted Average} - \text{Initial Score}}{\text{Initial Score}} \times 100$$

3. Drop Penalty

We also add an extra drop penalty that penalizes an arbitrary symptom/metric score if it is less than the immediate previous score.

$$\text{Drop Penalty Factor} = \alpha \times \frac{\sum_{t=1}^{n-1} \max(0, \text{Score at } t - \text{Score at } t+1)}{15 \times (n-1)} \times 100$$

α (default 1.0), Adjusts the weight of the penalty relative to progress percentage. This ensures only negative differences (regressions) are penalized. If **Score at t+1 > Score at t**, the difference is negative, and max(0, negative number) evaluates to 0, so no penalty is applied. The opposite is true if **Score at t+1 < Score at t**.

There is also no over-penalization:

- If scores improve or remain stable, the penalty remains 0.
- Only actual regressions (e.g., relapses or dips in progress) are penalized.

4. Final Progress Score

The final progress score is then a difference between the progress percentage and the drop penalty factor.

$$\text{Final Progress Score} = \text{Progress Percentage} - \text{Drop Penalty Factor}$$

These first 4 parts are handled by the compute_progress(). To stabilize this value, we set a minimum value of 0 to prevent negative values.

Cumulative Scoring

A cumulative score is calculated by `compute_cumulative_score()`. Again, given the symptoms, theme metrics and overall sentiment from all sessions, we make use of weighted average to derive the final cumulative score. By default, it is calculated by combining:

- Symptom scores (70% weight).
- Theme scores (20% weight).
- Sentiment score (10% weight).

$$\text{Cumulative Score} = \frac{0.7 \cdot \text{Total Symptom Score} + 0.2 \cdot \text{Total Theme Score} + 0.1 \cdot \text{Sentiment Score}}{\text{Total Weights}}$$

However, these weights can be adjusted as needed in the function. For stability, the final score is conditioned between 0 and 100%.

Summarization and Recommendations

The implementation (`summarize_recommend()`) also generates:

- **Summary**: Highlights the patient’s overall progress based on all metric scores.
- **Recommendations**: Provides actionable insights to improve the patient’s condition.

Visualization

Time-series plots are generated using `matplotlib` in the `visualize_symptoms` function. Each symptom or theme (insight, nutrition etc) is plotted, with shorter score sequences padded to align with longer sequences. The images are encoded in Base64 and returned to the frontend for display.

Key Features

1. **Dynamic Symptom Scoring**: Automatically adapts to various therapy session notes.

2. **Progress Tracking**: Accounts for improvements, regressions, and overall trends.

3. **Comprehensive Insights**: Combines numerical scoring and AI-generated summaries.

4. **User-Friendly visualization**: Provides a clear view of progress through visualizations and concise summaries.

Conclusion

In summary, my implementation provides a structured and efficient approach to tracking patient progress, combining AI capabilities with a logical and mathematical scoring system to score and track progress of patients across multiple sessions. It also ensures therapists have actionable insights as reminders for better decision-making and patient care.

—

Challenges

- Since scoring is done by LLM, scores might be subjective, inconsistent or inaccurate if the same sessions are passed through the LLM.
- Session data might be incomplete and contain sensitive information.
- Scoring system may not be the best as it relies heavily on subjective interpretations of symptom severity.
- Processing multiple files with large datasets could lead to performance bottlenecks.

–

Future Considerations

- Engineer better prompts to ground the model and mitigate inconsistent scoring.
- Database integration to save tracked and scored symptoms and metrics to reduce api calls and compute efficiency.
- Incorporate predictive analytics to forecast patient outcomes based on current progress and track symptoms and behaviour.
- Gamify solution for better patient engagement.
- Include Engagement level into the solution as it gives information about patient withdrawal or belief in the therapy and associated treatments.