MindWatch: AI Blog

https://www.youtube.com/watch?v=IurT4dr59Qg

Riccardo Dennis William
Department of Information System
Hanyang University
Seoul, South Korea
riccaden@students.zhaw.ch

Meyer Yves
Department of Information System
Hanyang University
Seoul, South Korea
meyeryve@students.zhaw.ch

Wu Yunjie
Department of Information System
Hanyang University
Seoul, South Korea
19855420423@163.com

Pereira Leandro
Department of Information System
Hanyang University
Seoul, South Korea
pereilea@students.zhaw.ch

Daniel Brody
Department of Information System
Hanyang University
Seoul, South Korea
brodydan@students.zhaw.ch

***Abstract*— The tech industry faces rising mental health challenges due to high job demands and pressure, yet workplace support remains reactive. We propose MindWatch, a web-based tool that leverages artificial intelligence (AI) to assess and monitor employee well-being. By analyzing standardized mental health assessments (PHQ-9, GAD-7, MBI, PSQ) alongside work-related data, MindWatch identifies early signs of stress, anxiety, and burnout. This proactive approach enables organizations to take informed actions to support their workforce. MindWatch aims to foster healthier, more sustainable work environments and reduce mental health challenges in the tech industry**

Keywords— Mental Health, Employee Well-being, Artificial Intelligence

I. Introduction (Motivation and Expected Outcomes

In recent years, the tech industry has experienced rapid growth and innovation, which has led to increased pressure on employees to meet tight deadlines, manage complex projects, and maintain high performance levels. While this growth has brought many opportunities, it has also come with a significant downside: the mental and emotional well-being of employees is often overlooked. Long hours, high expectations, and constant demands have contributed to an alarming rise in stress, anxiety, and burnout within the tech sector. These mental health challenges are compounded by the fast-paced and competitive environment in which tech professionals work, making it difficult for employees to seek help or take necessary breaks without fear of repercussions or falling behind.

Despite growing awareness of mental health issues, workplace support mechanisms in many organizations are still largely reactive, addressing mental health concerns only after they become significant problems. This approach often fails to prevent employees from reaching a breaking point, leading to high turnover rates, decreased productivity, and overall dissatisfaction. In response to these challenges, there is a clear need for proactive measures that can help detect mental health concerns early and provide employees with the support they need before reaching a crisis point.

Our project, MindWatch, seeks to bridge this gap by developing a web-based tool designed specifically to assess and monitor the mental health of employees in the tech industry. The motivation behind MindWatch stems from the urgent need to create a system that not only identifies potential mental health concerns early but also empowers organizations to take informed, proactive actions to support their workforce in real-time. By leveraging cutting-edge technologies like artificial intelligence (AI) and integrating standardized mental health assessments, we aim to provide a comprehensive tool that will help employers better understand and manage their employees' mental health.

MindWatch will incorporate well-established mental health assessment tools such as the PHQ-9 (Patient Health Questionnaire), GAD-7 (Generalized Anxiety Disorder Scale), MBI (Maslach Burnout Inventory), and PSQ (Perceived Stress Questionnaire), which are widely used to evaluate depression, anxiety, stress, burnout, and other mental health indicators. These assessments will be supplemented by data derived from employees' work-related activities, such as repository changes, project statuses, deadlines, and other performance metrics. By combining these two sets of data—mental health assessments and work performance indicators—MindWatch will offer a more holistic understanding of an employee's well-being, helping organizations identify individuals at risk of burnout or other mental health issues before they escalate.

At the conclusion of this project, we aim to deliver a functional, AI-powered tool that offers insightful, data-driven feedback to organizations. This platform will not only help businesses monitor and assess employee well-being but also equip them with the insights necessary to foster healthier, more sustainable working environments. Ultimately, our goal is to contribute to a positive shift in the tech industry by reducing mental health challenges and creating workplaces where employees feel supported, valued, and empowered. By bridging the gap between awareness and proactive action, MindWatch aspires to be an essential tool in the movement towards better mental health practices in the workplace.

II. Dataset

A. Selecting a Dataset

The dataset used in this project comes from a collection available on Kaggle, titled Sentiment Analysis for Mental Health. It is a comprehensive compilation of mental health-related statements from various sources, aimed at developing intelligent mental health chatbots and performing sentiment analysis.

B. *Description of the Dataset*

The dataset is an amalgamation of raw data from multiple Kaggle sources, which have been cleaned and curated for the purpose of mental health analysis. It includes statements tagged with seven distinct mental health statuses, making it suitable for a wide range of applications, such as predicting mental health conditions and building chatbots. The dataset includes:

- 3k Conversations Dataset for Chatbot
- Depression Reddit Cleaned
- Human Stress Prediction
- Predicting Anxiety in Mental Health Data
- Mental Health Dataset Bipolar
- Reddit Mental Health Data
- Students Anxiety and Depression Dataset
- Suicidal Mental Health Dataset
- Suicidal Tweet Detection Dataset

C. *Mental Health Status Tags*

The dataset contains statements tagged with one of the following mental health statuses:

1. Normal
2. Depression
3. Suicidal
4. Anxiety
5. Stress
6. Bi-Polar
7. Personality Disorder

Each entry in the dataset represents a statement that has been carefully labeled according to the mental health condition it corresponds to. This is an invaluable resource for training machine learning models that aim to classify and analyze mental health responses.

D. *Data Collection*

The data has been collected from a diverse set of platforms, including Reddit, Twitter, and other mental health-related sources. The statements are reflective of various mental health states and are tagged accordingly, making the dataset ideal for:

Developing intelligent mental health chatbots: The dataset serves as training data for chatbot applications aimed at providing mental health support.

Sentiment analysis: It can be used to analyze the emotional tone and sentiment of the text, helping identify potential mental health concerns.

Research and mental health trend analysis: Researchers can use the data to study patterns and trends related to mental health in various populations.

This dataset's richness and diversity make it an excellent foundation for our platform, allowing us to build models that can accurately predict mental health statuses from text responses.

III. Methodology

For this project, we used **Logistic Regression**, a widely-used classification algorithm, to predict the mental health status of employees based on text data. Logistic Regression is an effective model for binary and multi-class classification tasks, and it is particularly well-suited for text classification problems, such as predicting emotions or mental health states from textual responses.

We chose **Logistic Regression** because of its simplicity, interpretability, and efficiency, particularly when working with high-dimensional data like text. Additionally, it performs well even with sparse datasets, which is a common characteristic of text data after vectorization. We applied **TF-IDF (Term Frequency-Inverse Document Frequency)** to transform the raw text into numerical features, which can then be processed by the logistic regression model. TF-IDF helps highlight important words while reducing the impact of commonly occurring words that do not carry significant meaning for classification.

The **MaxAbsScaler** was used to scale the data, particularly for sparse matrices, as it is efficient and avoids unnecessary memory usage. The scaling ensures that all features contribute proportionately to the model's predictions, which is important for models like Logistic Regression that are sensitive to feature scaling.

1. Data Loading and Preprocessing:

The dataset is loaded from a CSV file containing the responses and their corresponding mental health statuses.

The preprocess_text function is applied to clean the raw text by converting it to lowercase, removing special characters, and eliminating extra spaces. This helps standardize the input text for further analysis.

```
def preprocess_text(text):
    text = text.lower()  # Convert to lowercase
    text = re.sub(r'\W', ' ', text)  # Remove special characters
    text = re.sub(r'\s+', ' ', text)  # Remove extra spaces
    return text.strip()  # Remove leading/trailing spaces
```

2. Data Cleaning:

Any rows with missing text data are dropped, ensuring that only valid, non-empty text is used for analysis.

3. TF-IDF Vectorization:

TF-IDF is used to convert the raw text into numerical data. This vectorization technique converts each document into a vector of numbers representing the importance of each word within the document relative to the entire dataset.

```
vectorizer = TfidfVectorizer(max_features=5000)  # Limit to 5000 most important features
X_train_tfidf = vectorizer.fit_transform(X_train)  # Fit and transform training data
X_test_tfidf = vectorizer.transform(X_test)  # Transform test data
```

4. Scaling with MaxAbsScaler:

MaxAbsScaler is used to scale the sparse matrix produced by the TF-IDF transformation. It scales the data by its maximum absolute value, making it more suitable for sparse data without introducing dense data representations.

5. Model Training:

A Logistic Regression model is trained using the transformed and scaled data. The max_iter=1000 parameter ensures that the model can converge to an optimal solution within 1000 iterations. The solver used is 'saga', which is particularly suited for large datasets and sparse matrices.

```
model = LogisticRegression(max_iter=1000, solver='saga', verbose=1)
model.fit(X_train_tfidf_scaled, y_train)  # Train the model once
```

6. Model Evaluation:

The trained model is then evaluated on the test set using classification_report, which provides metrics such as precision, recall, F1-score, and accuracy. These metrics are crucial for assessing the model's performance in predicting the correct mental health status.

```
predictions = model.predict(X_test_tfidf_scaled)  # Predict on test data
print(classification_report(y_test, predictions))  # Display performance metrics
```

7. Saving the Model and Vectorizer:

Once the model is trained, it is saved using joblib, which allows for easy loading of the model and vectorizer in future sessions without the need for retraining.

```
joblib.dump(model, 'emotion_model.pkl')  # Save the trained model
joblib.dump(vectorizer, 'tfidf_vectorizer.pkl')  # Save the TF-IDF vectorizer
```

IV. Evaluation & Analysis

The Logistic Regression model was trained and evaluated on the dataset, producing a classification report with the following performance metrics: precision, recall, and F1-score for each mental health category. The overall accuracy of the model is 76%.

Classification Report Summary

Anxiety: The model achieved a precision of 0.81, recall of 0.79, and F1-score of 0.80 for this class, which indicates good predictive performance for anxiety-related responses.

Bipolar: Precision is 0.86, while recall is 0.69, resulting in an F1-score of 0.77. While precision is high, the recall value suggests that the model may miss some bipolar cases.

Depression: The model scored 0.68 precision, 0.72 recall, and 0.70 F1-score for depression, indicating moderate performance in predicting this category.

Normal: The model performed best on this category, with a precision of 0.86, recall of 0.94, and F1-score of 0.90, suggesting that the model is highly effective in identifying normal cases.

Personality Disorder: Precision is 0.86, but recall is significantly lower at 0.49, resulting in an F1-score of 0.63. This suggests that personality disorder cases are harder to detect.

Stress: With a precision of 0.69 and recall of 0.46, the model has a lower F1-score of 0.55 for stress, indicating that stress cases are often missed or misclassified.

Suicidal: The model performed with 0.68 precision, 0.66 recall, and an F1-score of 0.67, reflecting a moderate ability to predict suicidal cases.

Overall Performance:

Accuracy: The overall accuracy of the model is 76%, which is decent, though the performance varies across different categories. The model performs particularly well in detecting "Normal" responses but struggles with categories such as "Stress" and "Personality Disorder."

Macro Average: The macro average, which treats all categories equally, shows a precision of 0.78, recall of 0.68, and F1-score of 0.72.

Weighted Average: The weighted average, which accounts for the class distribution, indicates an overall F1-score of 0.76, reflecting the model's balanced performance across all classes.

Analysis

The model's performance is generally strong for categories with a larger sample size, such as "Normal" and "Depression," but weaker for less represented classes, such as "Personality Disorder" and "Stress."

The low recall in certain categories (e.g., "Personality Disorder" and "Stress") suggests that the model may miss many cases in these categories, which is critical for a mental health analysis tool. To address this, future iterations could involve techniques like class balancing, oversampling, or using different model architectures (e.g., neural networks or ensemble methods) to improve performance for these underrepresented classes.

Conclusion

The Logistic Regression model provides a reasonable foundation for predicting mental health conditions based on text data, but further improvements are needed to enhance detection for certain categories, particularly "Stress" and "Personality Disorder." Future work could involve improving the model's recall and handling class imbalance, as well as exploring more advanced models for better performance in mental health classification.bar.

```
Training Logistic Regression model...
convergence after 25 epochs took 1 seconds

Classification Report:
                      precision    recall  f1-score   support

             Anxiety       0.81      0.79      0.80       767
             Bipolar       0.86      0.69      0.77       562
          Depression       0.68      0.72      0.70      3011
              Normal       0.86      0.94      0.90      3321
Personality disorder       0.86      0.49      0.63       217
              Stress       0.69      0.46      0.55       524
            Suicidal       0.68      0.66      0.67      2105

            accuracy                           0.76     10507
           macro avg       0.78      0.68      0.72     10507
        weighted avg       0.76      0.76      0.76     10507
```

V. RELATED WORK

Several studies and tools have laid the groundwork for this project, highlighting the potential of combining machine learning with mental health assessments to create actionable insights. Below are some key references and tools that have informed our approach:

1. Sentiment Analysis in Mental Health: *Authors and Affiliations*

Existing research has demonstrated the effectiveness of sentiment analysis in detecting mental health issues, especially when combined with structured questionnaire data. Thombs et al. (2014) explored the diagnostic accuracy of the PHQ-9 and other similar tools for detecting depression. This research shows how sentiment analysis can complement structured questionnaires like the GAD-7 and Maslach Burnout Inventory (MBI) to enhance the overall understanding of an individual's mental health. By analyzing free-text responses, sentiment analysis can capture nuanced feelings that standard questionnaire responses might miss.

2. Key Questions for Our Tool

For our tool, we have created five key combined questions to assess employee mental health. These questions are designed to evaluate various dimensions of mental well-being, such as motivation, stress, burnout, task management, and emotional burdens. By analyzing both the structured responses and sentiment analysis of free-text answers, our tool can provide valuable insights into the emotional and psychological states of employees. Here are the five key questions:

"How have you been feeling recently regarding your motivation and energy levels at work?"

This question helps gauge the employee's level of motivation and energy, which is critical for identifying early signs of burnout or disengagement. It is a combination of elements from Maslach Burnout Inventory (MBI) and PHQ-9 (regarding lack of interest and energy).

"Do you experience moments where you feel particularly stressed or overwhelmed? What causes this feeling?"

This question addresses stress and anxiety, drawing on elements from GAD-7 (Generalized Anxiety Disorder Scale) and the PSQ (Perceived Stress Questionnaire). It allows the tool to identify specific work-related stressors and the intensity of those stressors.

"How do you handle the pressure and demands of your job? Have you felt burnt out lately?"

This question combines aspects of Maslach Burnout Inventory (MBI) (specifically emotional exhaustion and burnout) and GAD-7 (relating to anxiety under pressure). It provides insight into how employees cope with stress and whether they feel overwhelmed by their job demands.

"Do you feel like you're able to manage your tasks effectively, or are there areas where you struggle?"

This question assesses an employee's perceived ability to manage workload and task-related stress, which is important for detecting potential signs of burnout or anxiety related to work. It combines elements from the MBI (Personal Accomplishment) and GAD-7 (nervousness and worry).

"Have you had moments recently where you felt particularly emotionally burdened? Would you like to share more?"

This question targets emotional exhaustion and emotional burden, which are key components of both PHQ-9 (depression) and MBI (burnout). It provides employees with the opportunity to elaborate on feelings of emotional strain, which the AI can analyze using sentiment analysis.

These five questions cover key areas of mental health, allowing the tool to detect early signs of stress, anxiety, burnout, and other psychological issues. By combining quantitative responses with qualitative insights, MindWatch can provide a comprehensive and nuanced assessment of an employee's mental well-being.

VI. CONLUSION

The development of MindWatch, a web-based tool designed to assess and monitor employee mental health, represents a significant step toward creating a more proactive approach to mental well-being in the workplace. As the tech industry continues to grow and evolve, the importance of addressing mental health issues among employees becomes increasingly critical. With high job demands, tight deadlines, and complex projects, professionals in the tech industry are particularly susceptible to stress, burnout, anxiety, and depression. The traditional reactive approaches to mental health support, where intervention occurs only after problems have escalated, are no longer sufficient. MindWatch aims to bridge this gap by providing a proactive, data-driven solution that can detect early signs of mental health issues and empower organizations to take action before problems become critical.

The tool's use of standardized mental health assessments combined with AI-driven analysis of free-text responses allows for a comprehensive understanding of an employee's well-being. By integrating both structured questionnaire data and sentiment analysis, MindWatch offers a more nuanced and accurate view of mental health, enabling companies to monitor employee well-being in real time and respond accordingly.

The performance of the Logistic Regression model, while solid with an overall accuracy of 76%, has highlighted areas for improvement. The model performs well for some categories, such as Normal and Anxiety, but struggles with others, especially those with less representation, like Stress and Personality Disorder. This imbalance in performance demonstrates the challenges of working with real-world, unbalanced datasets and the importance of using techniques like class balancing or exploring other machine learning models (such as Random Forests or Neural Networks) to improve predictions for underrepresented classes.

Additionally, the choice of features and the combination of structured survey questions with free-text responses enhances the model's ability to understand the complexity of mental health issues. For example, questions like, "How do you handle the pressure and demands of your job? Have you felt burnt out lately?" provide valuable insights into employee coping mechanisms, which are crucial for detecting burnout and stress early on. The AI model processes these insights by using Natural Language Processing (NLP) techniques, such as sentiment analysis and topic modeling, to extract emotional tone and detect patterns that might be missed through structured responses alone.

While the current version of MindWatch offers a solid foundation for early mental health detection, there are areas for future development:

Improvement in Model Performance: As we collect more data and refine the model, we aim to address the issues of class imbalance and improve the recall for categories like Stress and Personality Disorder, where the model's performance has been weaker.

Integration with Other Employee Data: In future iterations, MindWatch could be expanded to integrate other data sources, such as performance metrics, attendance records, and feedback from managers, to further personalize the mental health assessments and provide even more targeted interventions.

Real-Time Feedback: One of the next steps will be implementing a real-time feedback system, where employees receive immediate insights based on their responses. This feature will provide not only organizations but also individual employees with actionable advice to improve their well-being.

Expansion to Other Industries: While this tool is currently designed for the tech industry, the methodology behind MindWatch can be adapted to various other sectors facing similar mental health challenges, including healthcare, education, and customer service.

In conclusion, MindWatch has the potential to revolutionize how companies approach employee well-being. By leveraging AI and comprehensive mental health assessments, it provides a proactive tool that helps organizations identify and address mental health issues before they escalate. The continued development of this tool will ensure that organizations can create a healthier, more supportive work environment, ultimately leading to more productive, engaged, and satisfied employees.