Questionnaire in User Study

1. Personal Information

1.1 How old are you?

- Below 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 or older

1.2 What is your highest level of education?

- High school or below
- Bachelor's degree
- Master's degree
- Doctoral degree

1.3 What is your gender?

- Male
- Female

1.4 How familiar are you with computer science?

- Very familiar
- Somewhat familiar
- Neutral
- Somewhat unfamiliar
- Not familiar at all

1.5 How familiar are you with deep learning?

- Very familiar
- Somewhat familiar
- Neutral
- Somewhat unfamiliar
- Not familiar at all

1.6 How much do you agree with this statement: "I trust artificial intelligence and automation"?

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

1.7 How would you rate the difficulty of the following tasks from easiest to hardest?

- Understanding health advice from family and friends
    - Very easy
    - Somewhat easy
    - Neutral
    - Somewhat difficult
    - Very difficult
- Understanding nutritional information on food packaging
    - Very easy
    - Somewhat easy
    - Neutral
    - Somewhat difficult
    - Very difficult
- Understanding health-related information in the media
    - Very easy
    - Somewhat easy
    - Neutral
    - Somewhat difficult
    - Very difficult
- Understanding information on maintaining physical and mental health
    - Very easy
    - Somewhat easy
    - Neutral
    - Somewhat difficult
    - Very difficult

2. Understanding Domain Knowledge

Before evaluating the interpretability of the prediction model, it is essential to understand the widely used diagnostic criteria for depression. Please read and familiarize yourself with the following information on depression diagnosis and answer the questions for a knowledge check.

Depression Diagnosis Criteria (ICD-11)

The symptom criteria for depression include three core symptoms: low mood, loss of interest or pleasure, and low energy or fatigue. It also includes seven additional symptoms: reduced attention, low self-esteem, guilt and feelings of worthlessness, pessimism, self-harm and suicidal ideation, sleep disturbances, and decreased appetite.

The duration criterion for depression is that symptoms must persist for two weeks or more.

A diagnosis of depression is made when at least one core symptom and two additional symptoms are present.

The greater the number of symptoms, and the more frequent and persistent they are, the more severe depression is considered.

After reading the above knowledge, please answer the following questions:

2.1 According to the diagnostic criteria for depression, a person must meet:

- At least 2 core symptoms and 2 additional symptoms
- At least 1 core symptom and 2 additional symptoms
- Any 5 symptoms

2.2 Which of the following is a core symptom of depression?

- Sleep disturbance
- Depressed mood
- Feelings of guilt and worthlessness

2.3 The duration criteria for depression are:

- Symptoms persist for at least 2 weeks and occur almost every day
- Symptoms persist for at least 1 month and occur 2–3 days per week
- Symptoms persist for at least 6 months and occur every day

2.4 Which of the following situations supports the diagnosis of depression?

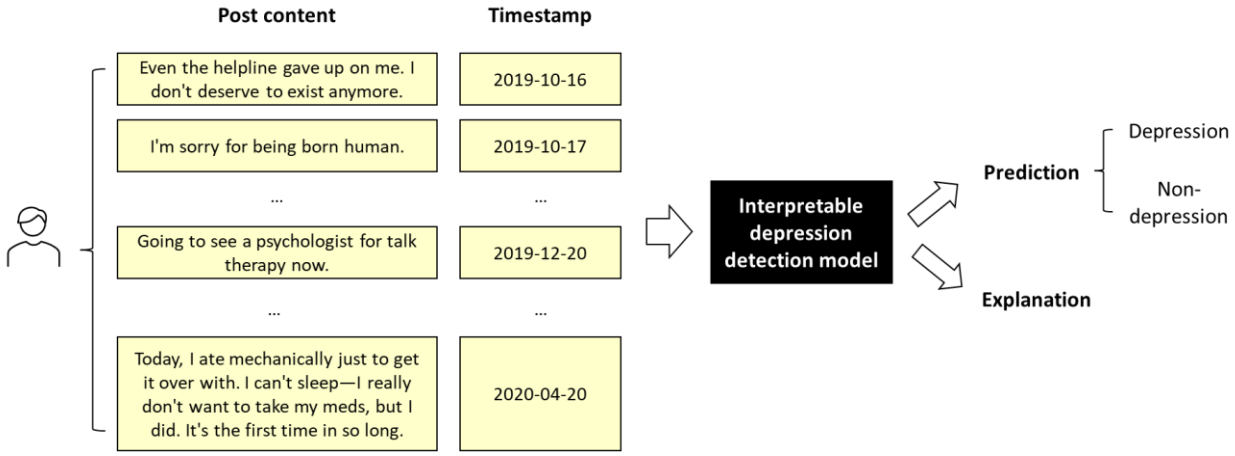
- Symptoms are present but do not significantly affect study, work, or daily life
- Symptoms cause significant distress and functional impairment in life
- Symptoms occur briefly only during setbacks and disappear completely afterward

2.5 Which of the following best reflects the overall diagnostic principle of depression?

- The more symptoms and the longer they persist, the more likely it is depression
- If depressed mood is present, it must be depression
- If sleep disturbance occurs, depression can be diagnosed

3. Task description

The original input and output results of the interpretable depression patient detection model are shown in the following figure (taking a real user as an example):



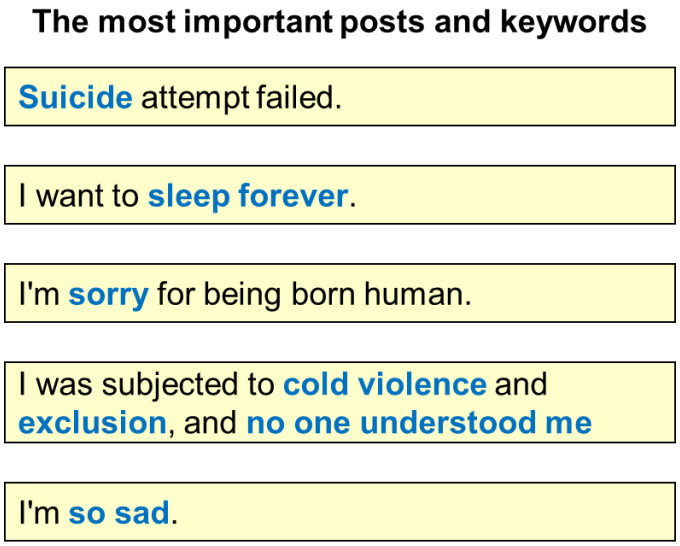
Next, for a user identified as having depression, we will randomly show you an explanation of a model and ask you to make a series of evaluations.

4. Model Explanation

Participants are randomly assigned to Group A, B, C, D, or E)

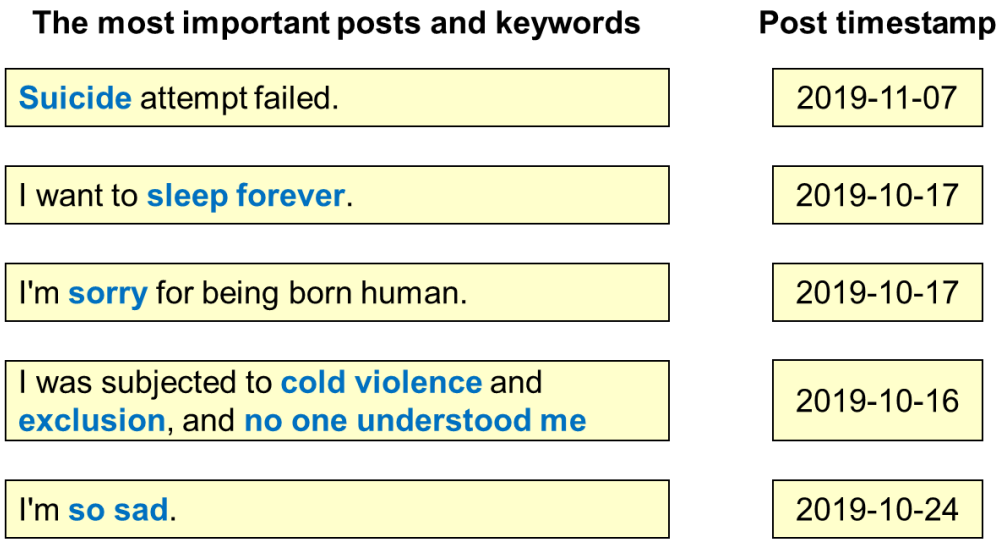
Group A

The model works by identifying posts related to depressive symptoms and their key words based on users' social media posts. The more related words there are, the higher the probability that the user is a patient with depression. The explanation provided by this model is as follows:



Group B

The model works by identifying posts related to depressive symptoms and their key words based on users' social media posts. The more related words there are, the higher the probability that the user is a patient with depression. The explanation provided by this model is as follows:



Group C

The model works by identifying the types of depressive symptoms based on the user's social media posts. The more depressive symptoms there are, the higher the probability that the patient is a sufferer of depression. The explanation provided by this model is as follows:

Depressive symptom	Posts corresponding to the symptoms
Suicidal thoughts	I want to die now.
Depressed mood	I'm in pain.
Fatigue	I'm so tired. I kind of want to give up.
Appetite Disorder	I have no appetite, no interest in food.
Self-blame	I'm sorry for being born human.

Group D

The model works by identifying the types of depressive symptoms based on the user's social media posts. The more depressive symptoms there are, the higher the probability that the patient is a sufferer of depression. The explanation provided by this model is as follows

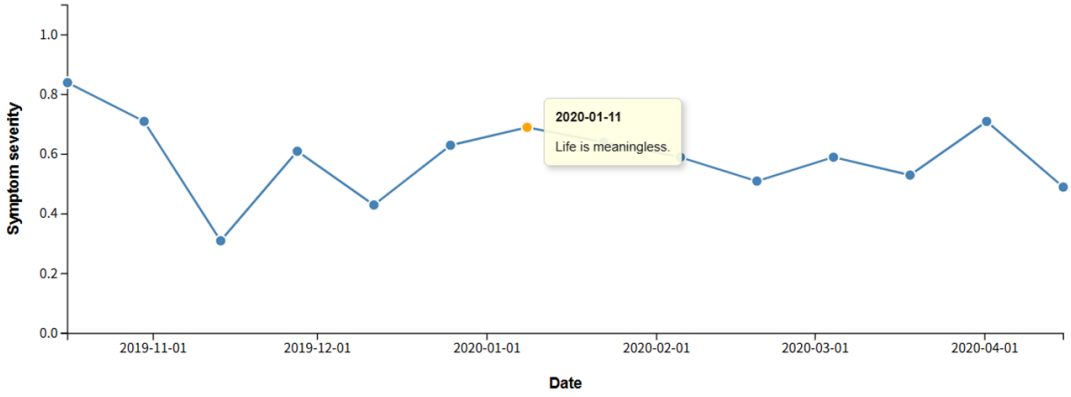
Depressive symptom	Posts corresponding to the symptoms	Post timestamp
Suicidal thoughts	I want to die now.	2019-10-18
Depressed mood	I'm in pain.	2019-10-16
Fatigue	I'm so tired. I kind of want to give up.	2019-10-17
Appetite Disorder	I have no appetite, no interest in food.	2019-10-18
Self-blame	I'm sorry for being born human.	2019-10-17

Group E

The model works as follows: First, identify the typical posts that disclose symptoms of depression in the target platform (i.e., example posts). Then, calculate the similarity between the user's posts and each of the typical symptom posts. Next, combine the posting time of the posts to infer the depression-related symptoms the user has and their duration. Finally, determine whether the user is suffering from depression. The explanation is as follows:

Depressed Mood

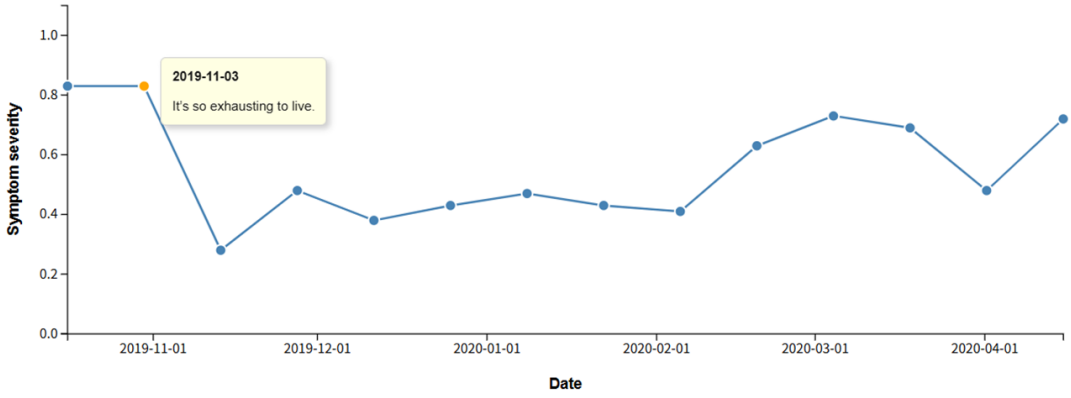
Example post: *Just existing, not really living. The world feels colorless and empty.*



Note: Hover over a point in the chart to view the corresponding symptom post.

Fatigue

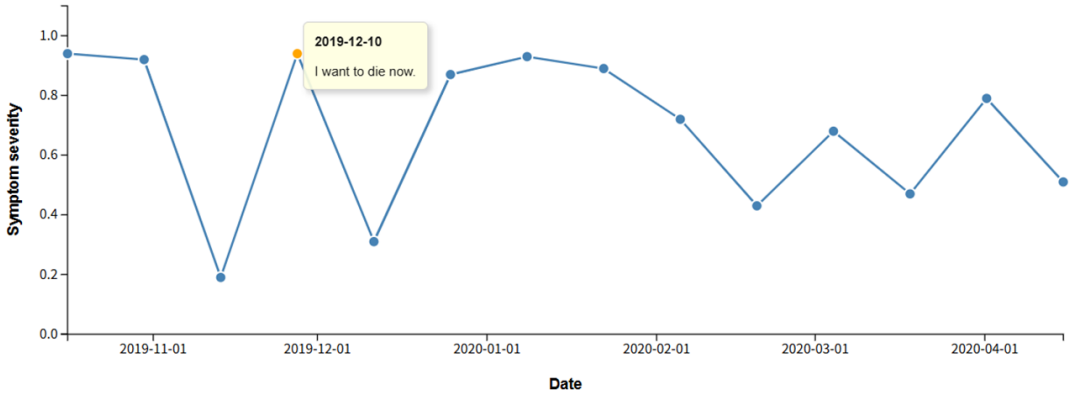
Example post: *Exhausted from life. Just being alive feels draining.*



Note: Hover over a point in the chart to view the corresponding symptom post.

Suicidal Thoughts

Example post: *Standing by the window, wanting to jump... I'm leaving tonight, I hope it won't hurt.*



Note: Hover over a point in the chart to view the corresponding symptom post.

5. Evaluation of Model Interpretability

5.1 Based on your evaluation of the understandability of your explanation of the model, please rate the following statements:

- I believe in the system's decision-making for individual patients.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
- I found the system complicated.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
- I found the system confusing.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
- I found the system easy to understand.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
- Please just select Neutral
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
- I found the system hard to remember.
    - Strongly disagree
    - Disagree

    - Neutral
    - Agree
    - Strongly agree
- I found the system predictable.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
- I understood how the system works.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
- I understood the relationship between inputs and output.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree

5.2 What aspects of the explanation helped you understand the model's decision?

5.3 What aspects of the explanation confused you in understanding the model's decision?

5.4 Do you have any other comments?
