



A histogram showing the distribution of token frequencies for two classes: 'suicide' (blue bars) and 'non-suicide' (orange bars). The x-axis is 'Token Frequency' on a logarithmic scale from  $10^0$  to  $10^3$ . The y-axis is 'Count' from 0 to 2500. The 'non-suicide' distribution is unimodal and peaks at a frequency of approximately 10 with a count of about 2500. The 'suicide' distribution is bimodal, with a smaller peak at frequency 10 (count ~600) and a larger peak at frequency 100 (count ~1900).

**Distance Map (multidimensional scaling)**

The distance map shows the relationship between different types of cases. The x-axis is labeled 'Type 1' and the y-axis is labeled 'Type 2'. The map shows a cluster of points for Type 1 and a cluster for Type 2, with a few points in between.

**Integration Statistics**

The Venn diagram shows the overlap between Type 1 and Type 2 cases. The x-axis is labeled 'Type 1' and the y-axis is labeled 'Type 2'. The overlap is labeled '21%'.

**Top-10 Most Relevant Terms for Type 1 (21% of cases)**

The bar chart shows the top 10 most relevant terms for Type 1 cases. The x-axis is labeled 'Frequency' and the y-axis is labeled 'Term'. The terms are: 'Type 1', 'Type 2', 'Type 3', 'Type 4', 'Type 5', 'Type 6', 'Type 7', 'Type 8', 'Type 9', and 'Type 10'. The frequency for each term is shown as a bar.

**Figure 1: Class frequency based on document length.**

The figure consists of two parts. The top part is a chord diagram showing relationships between 20 categories of documents. The categories are arranged in a circle and include: 'Depression', 'Suicide', 'Mental health', 'Mental illness', 'Mental health services', 'Mental health care', 'Mental health professionals', 'Mental health research', 'Mental health policy', 'Mental health law', 'Mental health ethics', 'Mental health education', 'Mental health training', 'Mental health practice', 'Mental health advocacy', 'Mental health activism', 'Mental health community', 'Mental health culture', 'Mental health history', and 'Mental health future'. The bottom part is a bar chart showing the frequency of documents for Depression (blue) and Suicide (orange) across different document lengths. The x-axis represents Document Length (0 to 1500) and the y-axis represents Frequency (10<sup>0</sup> to 10<sup>5</sup>).

Document Length	Depression Frequency	Suicide Frequency
0	~100,000	~100,000
100	~10,000	~10,000
200	~1,000	~1,000
300	~100	~100
400	~10	~10
500	~1	~1
600	~0.1	~0.1
700	~0.01	~0.01
800	~0.001	~0.001
900	~0.0001	~0.0001
1000	~0.00001	~0.00001
1100	~0.000001	~0.000001
1200	~0.0000001	~0.0000001
1300	~0.00000001	~0.00000001
1400	~0.000000001	~0.000000001
1500	~0.0000000001	~0.0000000001

■ suicide  
■ non-suicide

Depression  
Suicidal Risk

