**Assessment of the Suicide Text Mining algorithm and**

**Classification Model**

**Assessment 1: Coverage of the suicide attempt text mining algorithm**

**Introduction**

This study assesses the coverage or the face validity assessment of the suicide attempt text mining algorithm put together and the

**Method to Assess Coverage**

Random 50 documents were selected from SQL (Suicide attempt GATE table).

The document IDs were used to peruse through the entire document on CRIS front end.

All the instances of suicide attempt mentioned in the document were recorded and then compared to the results of the app.

**Results**

**To be recorded.**

**Assessment 2: Precision and Recall of suicide attempt classification model**

**Introduction**

This study uses Text Hunter a text mining and classification software that uses the text mining algorithm to identify phrases with a mention of suicide and then processes them for classification (manual annotation of the training set) into a positive or negative suicide attempt case. The classification tools used are primarily SVMs among other tools.

The models generated are assessed by comparing to a gold standard set of documents, annotated within Text Hunter by a clinician or researcher.

**Method to assess Precision and recall of Classification Model**

500 instances were annotated as gold standard. This was done after a high inter-rater agreement was achieved.

500 instances were annotated to training the classification model. The resultant model was assessed by comparing the model’s annotations of the gold standard with the manual annotations of the gold standard.

**Results**

Precision: 86%

Recall: 97%