TEXAS
The University of Texas at Austin

AI in HealthCare (AI 395T)

Professor Ying Ding

High-risk Project

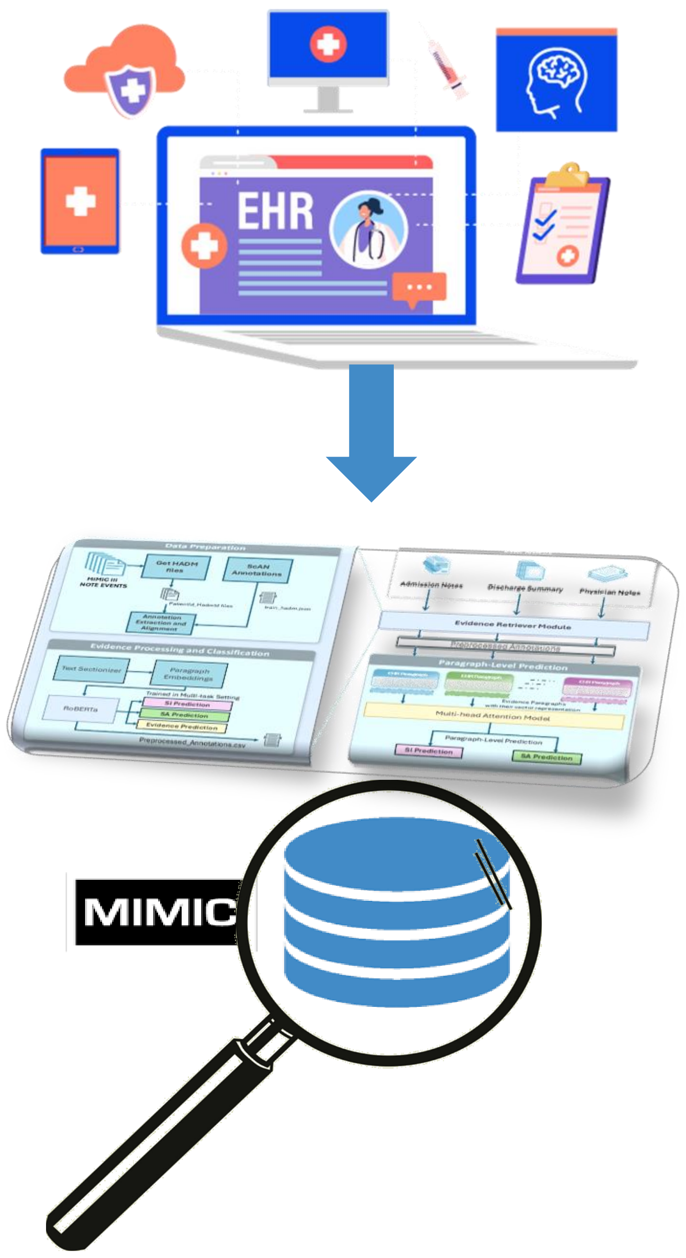
Detecting Suicidal Ideation and Suicide Attempts in Clinical Notes

An Adaptive Evidence Retrieval and Prediction Approach

Igor Carreon (IC9988)

Detecting Suicidal Ideation and Suicide Attempts in Clinical Notes

Table of Contents:



1. Background and Motivation: Adapting AI to Mental Health Analytics

Suicide is a pressing global health issue and a leading cause of preventable deaths. Behaviors such as suicidal ideation (SI) and suicide attempts (SA) are critical predictors of suicide risk. These behaviors are often recorded in narrative sections of electronic health records (EHRs), yet they remain challenging to analyze due to their unstructured nature. Traditional diagnostic systems, such as ICD codes, fail to capture these nuances, leading to underreporting and missed opportunities for timely intervention.

Advancements in natural language processing (NLP) provide a transformative opportunity to analyze unstructured EHR data and identify SI and SA more effectively. In this project, inspired by the ScANER framework, I developed a pipeline leveraging transformer-based models to detect and classify SI and SA within unstructured clinical notes.

The system consists of two primary components:

- The Evidence Retriever Module, which identifies paragraphs containing evidence of SI or SA.
- The Prediction Module, which categorizes these paragraphs into one of seven classes representing different levels of suicidality.

This project was particularly high-risk due to its complexity, the lack of publicly available note-level ground truth data, and the challenge of working independently to handle intricate data preprocessing, annotation, and model evaluation.

Extensive preprocessing included text segmentation, annotation alignment, and noise injection to improve model adaptability. The current system achieves paragraph-level classifications, but future work aims to extend these predictions to the clinical note level using strategies such as majority voting and temporal modeling. This project demonstrates the potential of AI to transform unstructured clinical text into actionable insights, paving the way for improved mental health evaluations and proactive suicide prevention strategies.

2. Related Work - Key Developments in SI and SA Detection

Rule-Based Approaches

- Early efforts (e.g., Downs et al., 2017) relied on manual rules
- Limitations: Lacked scalability and flexibility

Traditional Machine Learning

- Techniques like Random Forests (Breiman, 2001) and Support Vector Machines used for SA/SI classification
- Challenges: Relied on structured inputs and small datasets

Neural Networks

- Bhat and Goldman-Mellor (2017): Used feed-forward neural networks for suicide risk prediction on private datasets
- Common Issue: Limited accessibility to large, diverse datasets

Deep Learning Advancements

- Transformer models (e.g., RoBERTa) enable improved analysis of unstructured clinical text
- ScANER Framework (Rawat et al., 2022)
  - Integrated retrieval and prediction modules
  - Depended on proprietary datasets

Adopting an Alternative Approach

- Builds upon prior work in SI and SA detection within EHRs.
- Leverages publicly available tools and the MIMIC-III dataset.
- Implements transformer-based models (e.g., RoBERTa) for unstructured text analysis.
- Inspired by ScANER framework but employs independent, adaptable methods.
- Focuses on replicable and accessible solutions for broader applicability.

3. MIMIC-III and ScAN Dataset Overview

MIMIC-III DataSet

- Contains detailed information on patient demographics, vital signs, laboratory tests, medications, and interventions, including comprehensive clinical notes.
- Provides a rich variety of real-world clinical notes, enhancing model exposure to diverse clinical language.

ScAN Dataset

- The ScAN dataset is an expert-annotated subset of the MIMIC-III database, focusing on suicidal behavior.
- Contains 19,690 annotations for suicide attempts (SA) and suicidal ideation (SI) across 12,759 EHR notes.
- Annotations include evidence of suicidality and relevant sentences that indicate SA or SI during a patient's hospital stay.
- Experts compiled evidence to determine the presence of SA or SI events, including methods used for SA.
- An example of how SA annotations are presented can be seen in Figure 1, also found on the referenced document.
- The dataset enables a comprehensive analysis of how suicidality is documented and can be used to train AI models for detection and analysis.
- Offers focused annotations for SI/SA, enabling precise classification and reasoning tasks

Reference: ScAN Dataset (NAACL 2022)

Context:
[*Name*] is a 22 year old male with a history of angry and impulsive behavior who is transferred from an outside hospital s/p Tylenol overdose. [*Name*] reports that he and his girlfriend broke up last Wednesday, and that he subsequently went on an alcohol and cocaine binge lasting from Thursday to Saturday. He has used alcohol and cocaine regularly in the past, but he denies having had a binge of this quantity or duration before. On Saturday night, [**Name**] told his father that he had tried to hang himself at a nearby park, but the rope had broken.

Annotations:
Two instances of suicide attempts are annotated in the paragraph.
1. [*Name*] ... overdose: Annotated for suicide attempt and is assigned 'unsure' category as there is no definite documentation that it is a suicide attempt.
2. On Sat ... broken: This is also annotated for suicide attempt and is assigned the category: 'T71' (asphyxiation, hanging).

Figure 1: An example of *positive* and *unsure* evidence annotations for SA in an EHR note.

4. ScAN Dataset Details

Overview

- **ScAN (Suicide Attempt and Ideation Events Dataset)** focuses on identifying suicidal ideation (SI) and suicide attempts (SA) in clinical notes.
- Derived from **MIMIC-III**, a publicly available de-identified EHR dataset (2001–2012, Beth Israel Deaconess Medical Center).

Data Collection

- 697 hospital stays, covering 669 unique patients.
- Extracted 12,759 notes, including nursing notes, physician notes, and discharge summaries.
- Relevant sections identified using MedSpaCy clinical_sectionizer

Annotation Process

- Notes annotated for SA and SI by a trained annotator under a senior physician's supervision, offering targeted insights to improve classification accuracy and reasoning robustness
- Sentence-level annotations include:
  - Positive or Negative SA/SI.
  - Unsure label when intent is unclear.

Data Statistics

- **ScAN (Suicide Attempt and Ideation Events Dataset)** focuses on identifying suicidal ideation (SI) and suicide attempts (SA) in clinical notes.
- Derived from **MIMIC-III**, a publicly available de-identified EHR dataset (2001–2012, Beth Israel Deaconess Medical Center).

Annotation Structure Example


```json
{
    "8189_149399": {
        "Instance_625215": {
            "annotation": [
                "279",
                "301",
                null
            ],
            "category": "T36-T50",
            "period": "current",
            "frequency": "single",
            "suicide_attempt": null
        },
        "Instance_625230": {
            "annotation": [
                "10696",
                "10760",
                null
            ],
            "status": "absent",
            "suicide_ideation": null
        }
    },
    ..
    "7259_193543": {
        "Instance_616558": {
            "annotation": [
                "329",
                "560",
                null
            ],
            "status": "present",
            "suicide_ideation": null
        },
        "Instance_616559": {
            "annotation": [
                "329",
                "560",
                null
            ],
            "category": "T36-T50",
            "period": "current",
            "frequency": "single",
            "suicide_attempt": null
        },
```

5. The Significance of Annotation Data

ScAN Annotations

Annotation Precision

- Incorporates start and end indices to locate evidence within clinical texts.

Categorization

- Distinguishes between SI (suicidal ideation) and SA (suicide attempts) with detailed attributes like method and frequency.

Section Segmentation

- Utilizes MedSpaCy's clinical sectionizer to target relevant sections like 'Family History' and 'Assessment and Plan.'

Preprocessed Annotations by the Evidence Retriever Model

This module, developed as part of this project, builds upon the annotation framework by introducing additional steps to better simulate real-world data complexities to generalize and effectively identify SI and SA evidence within unstructured clinical text.

Noise Handling

- Introduces unrelated paragraphs to improve model robustness and simulate real-world scenarios.

Tokenization and Alignment

- Employs RoBERTa tokenizer and aligns annotations for SA and SI categories to provide clear training signals.

Dataset Focus

- Centers on notes tied to hospital admissions with ICD codes for suicide or overdose, capturing nuanced and comprehensive details.

6. Methodology - Overview

A Structured Approach for Suicide Risk Analysis

Objective

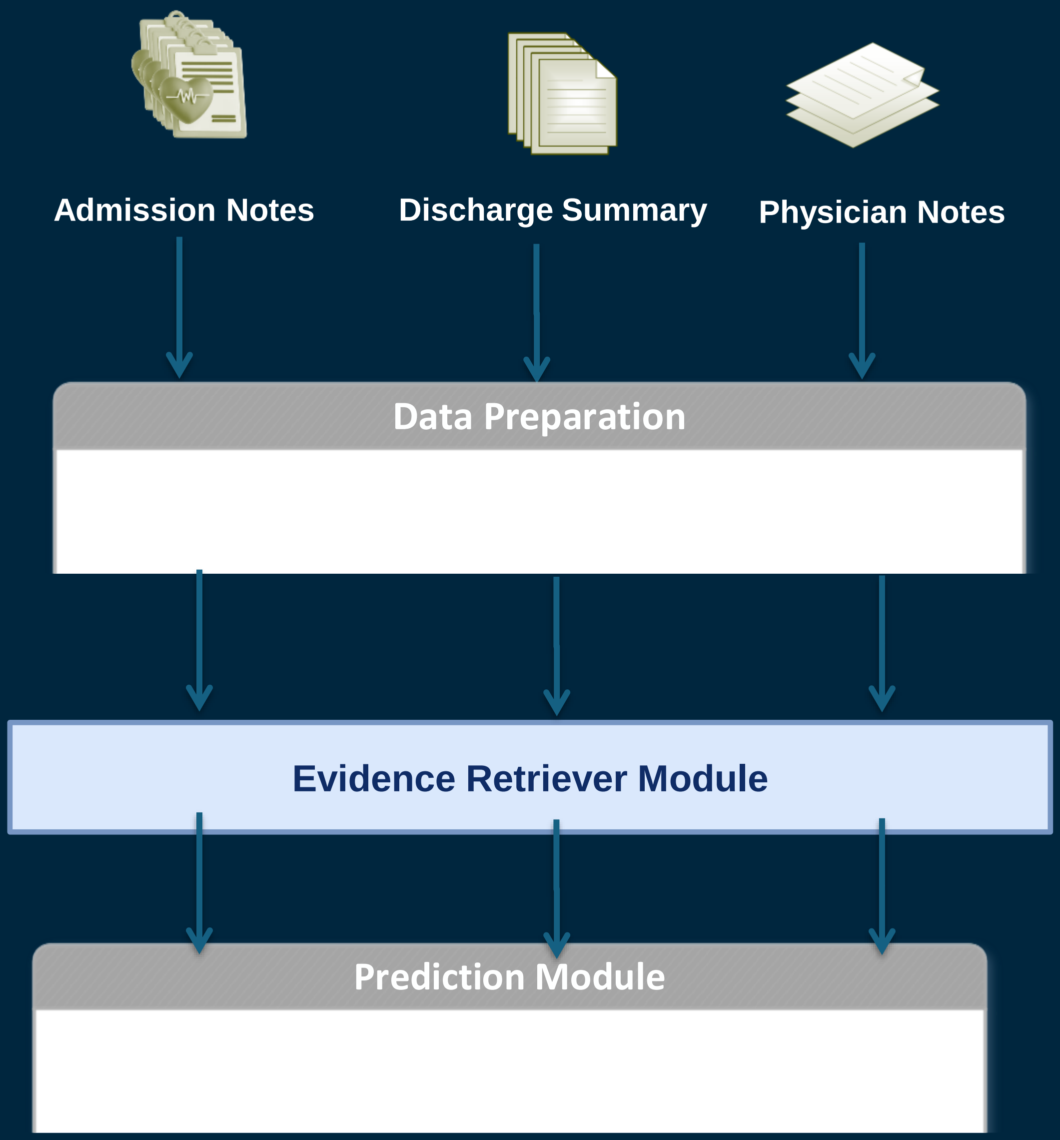
- Identify and classify suicidal ideation (SI) and suicide attempts (SA) within unstructured clinical text

Components

1. Data Preparation: Organizing and segmenting clinical notes.
2. Evidence Retriever Module: Locating relevant paragraphs containing evidence of SI and SA.
3. Prediction Module: Classifying identified paragraphs into predefined categories.

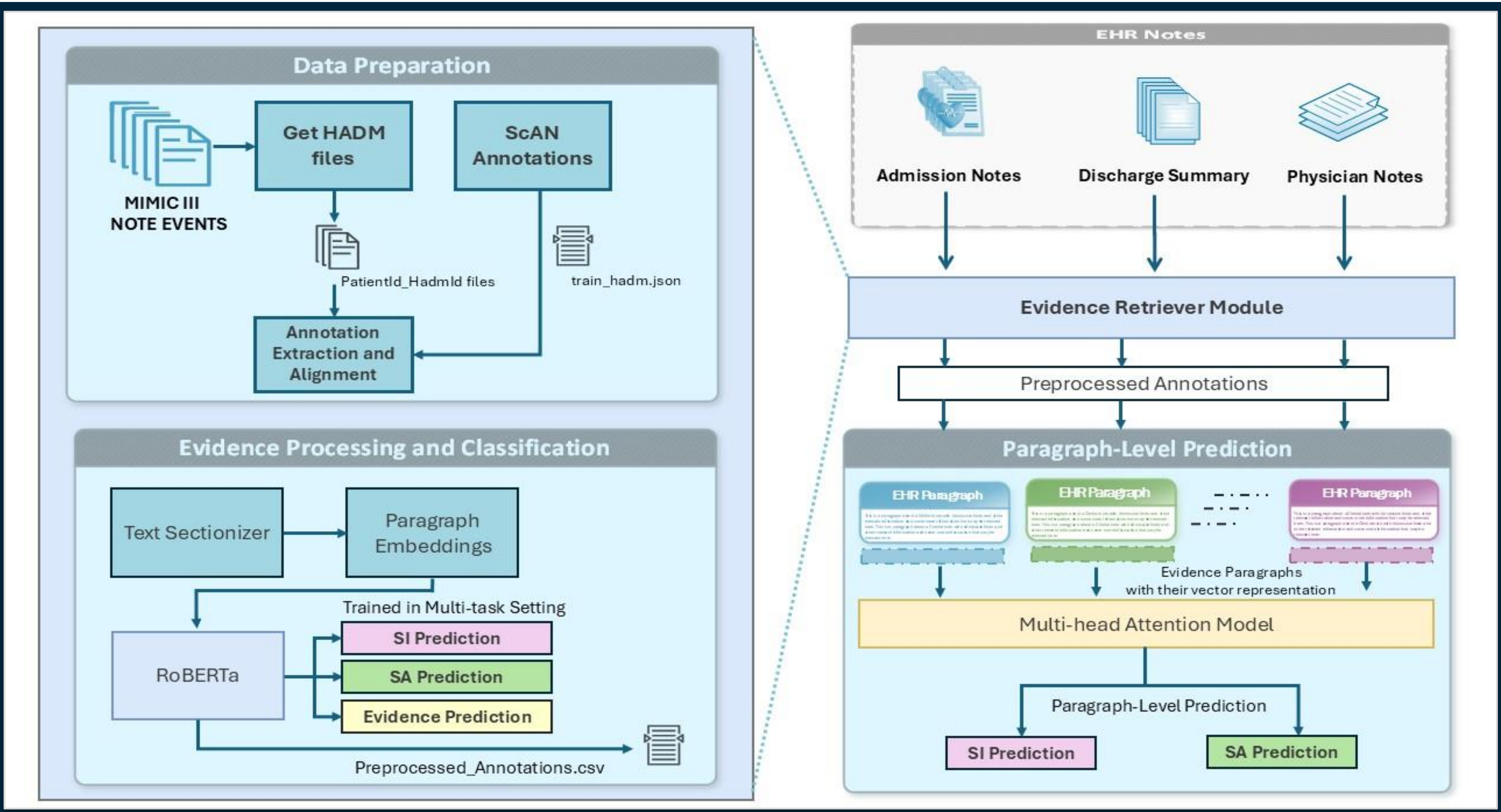
Key Techniques

- Data augmentation
- Transformer-based embeddings
- Attention mechanisms



7. Methodology - Workflow Diagram

Workflow for Evidence Retrieval and Paragraph-Prediction Level Modules



8. Methodology - Data Preparation

Data Acquisition and Annotation Preprocessing

Data Acquisition and Annotation Dataset

1. Dataset: Clinical notes extracted from the MIMIC-III database, focusing on categories most likely to contain SI and SA evidence (e.g., discharge summaries, physician notes).Selection Criteria: Notes linked to hospital admissions with ICD codes for suicide or overdose, ensuring high relevance to SI/SA indicators.
2. Annotation Precision: Expert annotations aligned with start and end indices, marking the exact text locations containing SI/SA evidence.

Annotation Preprocessing

1. Categorization: Each paragraph instance was categorized with one of the following labels: SA_Positive, SA_Negative, SI_Positive, SI_Negative, or Neutral-SA. The corresponding sentence was then extracted from the clinical notes to align with the assigned label.
2. Noise Handling: Introduced irrelevant paragraphs to simulate real-world conditions, improving the model's ability to differentiate between evidence and noise.

Output

A robust dataset enriched with structured evidence and realistic noise, forming a reliable foundation for the Evidence Retriever and Prediction Modules

9. Methodology - Evidence Retriever Module

Identifying Relevant Paragraphs

The Evidence Retriever Module identifies and classifies evidence of suicidal ideation (SI) and suicide attempts (SA) from clinical text. Using a fine-tuned RoBERTa model, it categorizes paragraphs into **evidence_yes** (relevant) or **evidence_no** (irrelevant), providing critical input for downstream predictive tasks. Beyond binary classification, it delivers detailed outputs across multiple SI/SA categories.

Key Features

1. Fine-Tuned Transformer: RoBERTa model optimized for evidence classification in clinical text.
2. Multi-Category Outputs: Classifies paragraphs into SI/SA categories for nuanced analysis.
3. Noise Handling: Improved generalization through the inclusion of irrelevant text.
4. Evaluation Metrics: Precision, recall, and F1-scores for comprehensive performance insights.

Output

1. Binary Evidence Labels: Each paragraph sentence classified as either evidence_yes or evidence_no.
2. Detailed Annotations: Outputs a structured file (processed_annotations.csv) with paragraph-level categories, including SA_Positive, SA_Unsure, SA_Negative, Neutral-SA, SI_Positive, SI_Negative Neutral-SI.
3. Foundation for Prediction Module: Ensures only relevant paragraphs are analyzed further, improving pipeline efficiency.

10. Methodology - Prediction Module

Detailed Classification for Suicidal Behaviors

The Prediction Module builds on the outputs of the Evidence Retriever to classify paragraphs into detailed categories of suicidal ideation (SI) and suicide attempts (SA). This module moves beyond binary evidence classification, providing nuanced insights into suicide-related behaviors within clinical notes.

Key Features

1. Multi-Label Classification: Assigns paragraphs to one of seven categories: SA_Positive, SA_Unsure, SA_Negative, Neutral-SA, SI_Positive, SI_Negative, Neutral-SI
2. Transformer-Based Architecture:
   - Leverages fine-tuned RoBERTa embeddings to capture paragraph context.
   - Implements a Multi-Head Attention mechanism for enhanced contextual understanding.
3. Weighted Sampling: Mitigates class imbalance by prioritizing underrepresented categories during training.
4. Iterative Training Strategy: Evaluated performance across multiple epochs to optimize accuracy and efficiency.

Output

1. Paragraph-Level Predictions: Each paragraph classified into one of the seven categories, enabling detailed analysis of SI/SA behaviors.
2. Structured Annotations for Analysis: Outputs include paragraph labels, confidence scores, and contextual embeddings to support downstream tasks.
3. Insights for Aggregation: Provides granular classifications that can be aggregated for patient-level predictions in future developments pipeline efficiency.

11. Results Overview - Evidence Retriever Module

Model Performance Metrics of Evidence Retriever Module

The Evidence Retriever Module was evaluated to determine its ability to identify and classify paragraphs containing evidence of suicidal ideation (SI) and suicide attempts (SA). Key findings from a representative run include:

SA_Positive: Achieved a high F1-score of 0.91, with precision at 0.89 and recall at 0.92, showcasing its ability to accurately identify explicit references to suicide attempts.
SA_Unsure: Moderate performance with an F1-score of 0.43, precision at 0.47, and recall at 0.39, indicating challenges in handling ambiguous cases.
SA_Negative: Limited success, with a low F1-score of 0.21, reflecting difficulty in identifying paragraphs without evidence of suicide attempts.
Neutral-SA: The model failed to recognize this category, achieving an F1-score of 0.00, pointing to issues like data imbalance or insufficient learning signals.
SI_Positive & SI_Negative: Moderate F1-scores of 0.61 and 0.65, respectively, highlighting strengths in detecting suicidality indicators but also areas needing improvement.
Neutral-SI: Performed well with a recall of 1.00, achieving an F1-score of 0.68, though precision suggests a tendency to overclassify as neutral.

Overall Performance
The module achieved an overall accuracy of 0.78. However, the macro average F1-score of 0.50 and weighted average F1-score of 0.75 point to disparities across categories, emphasizing the need for targeted refinements.

While the Evidence Retriever Module demonstrates robust performance in dominant categories like SA_Positive, it requires further development to address challenges in underrepresented categories. These refinements will ensure a more balanced and effective classification pipeline.

Label	*Precision*	*Recall*	*F1-Score*	*Support*
SA_Positive	0.89	0.92	0.91	1830
SA_Unsure	0.47	0.39	0.43	234
SA_Negative	0.4	0.14	0.21	29
Neutral-SA	0	0	0	204
SI_Positive	0.61	0.61	0.61	132
SI_Negative	0.67	0.64	0.65	81
Neutral-SI	0.51	1	0.68	212
Accuracy			0.78	2722
Macro Avg	0.51	0.53	0.5	2722
Weighted Avg	0.73	0.78	0.75	2722

12. Evidence Retriever Module Output

Structured Annotations and Classification Results

subject_id	hadm_id	instance_id	start_idx	end_idx	text_span	category	period	frequency	suicide_attemp	SA_CATEGORY	SI_CATEGORY	tokens	evidence
12259	123212	Instance_555912	276	289	Drug overdose	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 41533, 11972, 2]	evidence_yes
12259	123212	Instance_555913	291	358	Pt attempted suicide with [**Doctor Last Name 18928**] and	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 510, 90, 3751, 4260, 19, 646, 12606, 41152, 1426, 10704, 29304, 2517, 12606, 742,	evidence_yes
12259	123212	Instance_555914	437	620	44 y/o man with recent diagnosis with P	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 3305, 1423, 73, 139, 313, 19, 485, 9726, 19, 221, 50140, 2]	evidence_yes
12259	123212	Instance_555915	4309	4409	Psychiatry was consulted in the [**Hospital Unit Name	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 44353, 11284, 1506, 21, 18998, 11, 5, 646, 12606, 725, 44206, 7545, 10704, 25758,	evidence_yes
12259	123212	Instance_555916	5601	5620	Overdose on opiates	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 10777, 35238, 15, 5963, 31821, 2]	evidence_yes
12259	123212	Instance_555917	5621	5636	Suicide attempt	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 20689, 15772, 2120, 2]	evidence_yes
12259	123212	Instance_555918	6667	6675	OVERDOSE	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 36933, 495, 7949, 2]	evidence_yes
8189	149399	Instance_625230	10696	10760	Pt was asked if she is suicidal at this time and she stated "N	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Negative	[0, 510, 90, 21, 553, 114, 79, 16, 23630, 23, 42, 86, 8, 79, 2305, 22, 13449, 113, 2]	evidence_yes
62920	145660	None	None	None	None	Unknown	Unknown	Unknown	Unknown	Neutral-SA	Neutral-SI	[]	evidence_yes
12559	167599	None	None	None	None	Unknown	Unknown	Unknown	Unknown	Neutral-SA	Neutral-SI	[]	evidence_yes
7259	193543	Instance_616557	230	242	TCA overdose	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 6078, 250, 11972, 2]	evidence_yes
7259	193543	Instance_616558	329	560	Mr. [**Known lastname 24158**] is a 41yo man with h/o HIV	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Positive	[0, 10980, 4, 646, 12606, 47186, 94, 13650, 706, 26758, 12606, 742, 16, 10, 3492, 983	evidence_yes
7259	193543	Instance_616560	1731	1892	Notably the patient had a recent hospitalization in [**3-	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Positive	[0, 7199, 4735, 5, 3186, 56, 10, 485, 1098, 1938, 11, 646, 12606, 246, 12, 1646, 12606,	evidence_yes
7259	193543	Instance_616561	2030	2063	depression with suicidal ideation	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Positive	[0, 17272, 21791, 19, 23630, 23432, 1258, 2]	evidence_yes
7259	193543	Instance_616562	4542	4710	41yo man with h/o HIV, depression with suicidal ideation,	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Positive	[0, 4006, 9839, 313, 19, 1368, 73, 139, 7947, 6, 6943, 19, 23630, 23432, 1258, 6, 5011	evidence_yes
7259	193543	Instance_616563	4542	4710	41yo man with h/o HIV, depression with suicidal ideation,	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 4006, 9839, 313, 19, 1368, 73, 139, 7947, 6, 6943, 19, 23630, 23432, 1258, 6, 5011	evidence_yes
7259	193543	Instance_616564	4716	4728	TCA overdose	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 6078, 250, 11972, 2]	evidence_yes
7259	193543	Instance_616565	4730	4848	Pt with ?EKG changes that were concerning for	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 510, 90, 19, 17487, 717, 530, 534, 1022, 14, 58, 8082, 13, 50118, 26251, 11972, 36,	evidence_yes
7259	193543	Instance_616566	5183	5255	ECG changes resolved and	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 3586, 534, 1022, 8179, 8, 50118, 23846, 56, 117, 617, 526, 3038, 9, 11972, 2]	evidence_yes
7259	193543	Instance_616567	5261	5274	SSRI overdose	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 8108, 6934, 11972, 2]	evidence_yes
7259	193543	Instance_616568	5276	5316	treatment was same as TCA overdose above	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 37558, 21, 276, 25, 255, 4054, 11972, 1065, 2]	evidence_yes
7259	193543	Instance_616569	6367	6400	Depression with suicidal ideation	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Positive	[0, 29774, 21791, 19, 23630, 23432, 1258, 2]	evidence_yes
7259	193543	Instance_616570	6650	6909	Of note, the patient's PCP	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Positive	[0, 10643, 1591, 6, 5, 3186, 18, 4985, 510, 50118, 10975, 12606, 31723, 466, 36, 4386	evidence_yes
7259	193543	Instance_616571	7632	7653	TCA and SSRI overdose	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 6078, 250, 8, 11643, 6934, 11972, 2]	evidence_yes
7259	193543	Instance_616576	12633	12871	REPORT RECEIVED FROM [**Name (NI) **] NURSE, PT ARRIVE	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Positive	[0, 4629, 20275, 14895, 717, 6372, 1691, 11974, 646, 12606, 31723, 36, 17640, 43, 13	evidence_yes
18087	178389	Instance_561647	927	1069	Denied any suicidal ideation but does endorse depression	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Negative	[0, 27267, 2550, 143, 23630, 23432, 1258, 53, 473, 18839, 6943, 8, 50118, 1073, 5224,	evidence_yes
18087	178389	Instance_561648	4650	4660	LITHIUM OD	T36-T50	current	single		SA_Positive	Neutral-SI	[0, 574, 27698, 43145, 20778, 2]	evidence_yes
22045	141239	Instance_565229	1153	1355	Pt was seen by psych,	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Negative	[0, 510, 90, 21, 450, 30, 21968, 6, 50118, 28422, 352, 56, 117, 595, 17982, 18, 36, 463,	evidence_yes
22045	141239	Instance_565230	1153	1355	Pt was seen by psych,	N/A	N/A	N/A		SA_Negative	Neutral-SI	[0, 510, 90, 21, 450, 30, 21968, 6, 50118, 28422, 352, 56, 117, 595, 17982, 18, 36, 463,	evidence_yes
22045	141239	Instance_565231	4813	4854	as above, denies any SI/denies current SI	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Negative	[0, 281, 1065, 6, 9118, 143, 17982, 73, 3898, 918, 595, 17982, 2]	evidence_yes
22045	141239	Instance_565232	2683	2721	denies SI, stable, requesting to leave	Unknown	Unknown	Unknown	Unknown	Neutral-SA	SI_Negative	[0, 3898, 918, 17982, 6, 4375, 6, 14030, 7, 989, 2]	evidence_yes
29394	142086	Instance_576717	409	600	32 yr old male with a history of schizophrenia, suicide	T14.91	past	N/A		SA_Positive	Neutral-SI	[0, 2881, 39750, 793, 2943, 19, 10, 750, 9, 32285, 6, 4260, 2120, 6, 50118, 463, 3931,	evidence_yes

The Evidence Retriever Module outputs structured annotations by aligning text spans from clinical notes with categories such as SA_Positive, Neutral-SA, SI_Positive, and others. Each paragraph is classified as evidence_yes or irrelevant, and corresponding tokens are generated for model input. This output, exemplified in the table, captures key information such as start and end indices, text context, and categorical labels. By transforming unstructured text into actionable data, this module forms the foundation for further analysis in the Prediction Module, enabling accurate classification of suicidal ideation (SI) and suicide attempts (SA).

13. Results Overview - Prediction Module

Model Performance Metrics of Prediction Module

The Prediction Module underwent extensive evaluation to assess its ability to classify suicidal ideation (SI) and suicide attempts (SA) accurately. Key findings from a notable iteration include the following:

SA_Positive: Achieved a high F1-score of 0.91, with precision at 0.85 and recall at 0.99, indicating strong accuracy in detecting clear cases of suicide attempts.

SA_Negative: Moderate performance with an F1-score of 0.50, reflecting success in identifying paragraphs without evidence of suicide attempts.

SA_Unsure: F1-score of 0.58, with recall at 0.45, signaling the need for enhanced sensitivity to ambiguous cases.

Neutral-SA: Struggled significantly, with a low F1-score of 0.05, highlighting challenges in recognizing neutral content.

SI_Positive: Demonstrated limited effectiveness with an F1-score of 0.06, suggesting difficulty in classifying suicidality indications reliably.

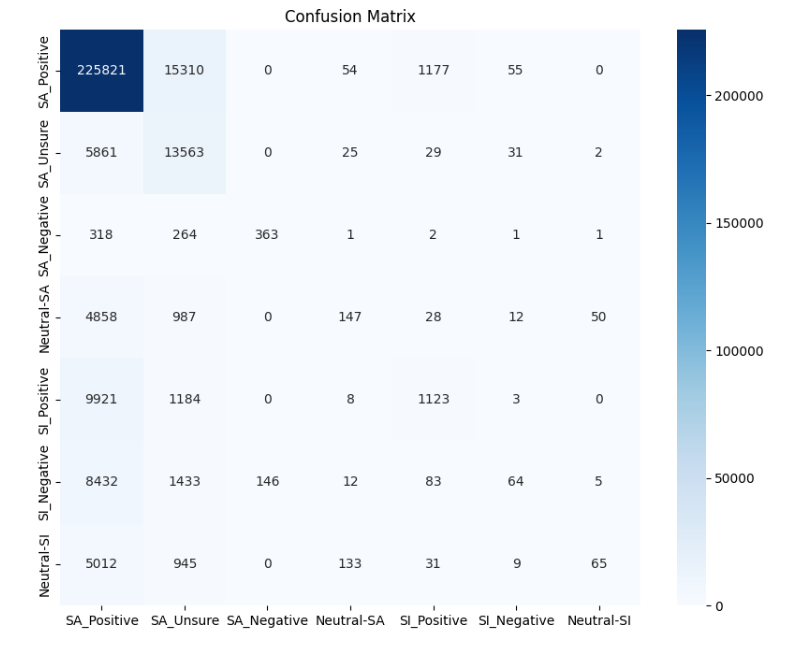
SI_Negative and Neutral-SI: Both categories showed low F1-scores of 0.01 and 0.08, respectively, indicating a need for improved classification of these categories.

Overall Performance

The module achieved an accuracy of 0.84, reflecting robustness for dominant classes. However, the macro average F1-score of 0.31 and weighted average F1-score of 0.79 reveal disparities in class-specific performance. These findings underscore the importance of refining the module to better balance training data and address underrepresented categories.

These results demonstrate the potential of the Prediction Module while emphasizing areas for future improvement. Continued development will enhance its ability to support mental health professionals in identifying and understanding suicide risks from unstructured clinical text.

Label	Precision	Recall	F1-Score	Support
SA_Positive	0.85	0.99	0.91	242417
SA_Unsure	0.82	0.45	0.58	19511
SA_Negative	0.71	0.38	0.5	950
Neutral-SA	0.36	0.02	0.05	6082
SI_Positive	0.38	0.03	0.06	12239
SI_Negative	0.25	0.01	0.01	10175
Neutral-SI	0.34	0.04	0.08	6195
Accuracy			0.84	297569
Macro Avg	0.53	0.28	0.31	297569
Weighted Avg	0.78	0.84	0.79	297569



14. Future Directions

Advancing from Paragraph-Level to Note-Level Insights

The pipeline demonstrates success in classifying paragraphs into suicide-related categories (e.g., SA Positive, SI Positive). However, to enhance clinical relevance, the next step involves aggregating paragraph-level predictions into unified clinical note-level classifications. This transition will provide a more holistic understanding of patient conditions.

Key Focus Areas

1. Note-Level Aggregator

Majority Voting: Assigns the most frequent paragraph label to the note. Initial testing showed limitations due to the lack of clinical note-level ground truth data.
Weighted Scoring: Assigns higher influence to predictions with greater confidence scores.
Temporal Modeling: Uses sequential models to analyze the progression of evidence across paragraphs, capturing context and narrative flow.

2. Ground Truth Development

Challenge: Absence of labeled clinical notes for note-level predictions.
Future Goals: Collaborate with clinical experts to create datasets with note-level labels, enabling validation and supervised training of the Note-Level Aggregator.

3. Additional Improvements

Neutral Categories: Explore aggregation strategies to ensure neutral classifications do not overshadow strong positive/negative evidence.
Scalability: Optimize modules to handle larger datasets or real-time clinical data.
Evaluation Metrics: Develop consistency metrics between paragraph and note-level predictions.
Explainability: Highlight influential paragraphs driving classifications to aid clinician interpretation.

These future directions aim to transition the pipeline into a fully functional system for clinical note-level classification. With improved scalability, explainability, and access to ground truth data, the project will move closer to real-world clinical implementation and impact.

15. Conclusion

Advancing AI for Mental Health Assessment

This project represents another step toward leveraging AI to address critical challenges in mental health assessment by analyzing unstructured clinical text for signs of suicidal ideation (SI) and suicide attempts (SA). By combining robust data preparation techniques, advanced NLP methodologies, and transformer-based models, the developed system demonstrates the potential to transform raw clinical notes into actionable insights for healthcare professionals.

Key Takeaways

- Developed an AI-driven pipeline to classify suicidal ideation (SI) and suicide attempts (SA) within unstructured clinical text.
- Evidence Retriever Module and Prediction Module showed promise in detecting clear instances of suicidal behavior but highlighted challenges in nuanced and ambiguous cases.

Future Vision

- Note-Level Aggregator Module: Transitioning from paragraph-level to comprehensive note-level interpretations.
- Ground Truth Data Acquisition: Enable validation and supervised training for clinical relevance.

Impact on Mental Health Assessment:

- Early and accurate identification of at-risk individuals.
- Support for timely interventions through enhanced AI-driven tools.
- Future efforts will focus on scalability, improved consistency, and explainability to realize AI's full potential in suicide prevention and healthcare diagnostics..