

SCORE-IT: A Machine Learning Framework for Automatic Standardization of EEG Reports

Samarth Rawal¹ and Yogatheesan Varatharajah²

¹ Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign

² Department of Bioengineering, University of Illinois at Urbana-Champaign

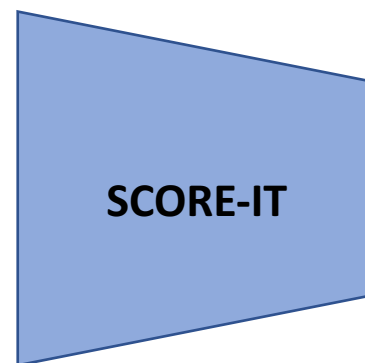
Introduction

- **Analysis and interpretation of electroencephalogram (EEG)** is important in the diagnosis of neurological conditions such as epilepsy.
- However, **lack of standardization** of EEG patient records makes automatic analysis challenging.
- **Standardized Computer-based Organized Reporting of EEG (SCORE)^[1] system:** collection of technical standards specifying terminology used and information included within reports.
- We propose a **machine learning-based system** that **automatically extracts components** from the SCORE specification from unstructured, natural-language EEG reports.

CLINICAL HISTORY: 26 year old right handed female with past medical history of epilepsy, has been seizure free for the past 7 years. This is a follow-up EEG...

DESCRIPTION OF THE RECORD: In wakefulness, the background EEG is appropriately organized with a 9-Hz posterior dominant rhythm. Hyperventilation produced supraharmonic driving. When...

Free-text EEG Report



Past Medical History: epilepsy

Findings: 3-5 Hz spike and wave activity, ...

Classification: ABNORMAL

Diagnostic Findings: "generalized mechanism of epilepsy"

Structured Data

Data

- Designed and validated on TUH EEG dataset^[5]: 16,000 EEG reports; classification of:
 1. whether the patient is being **evaluated for epilepsy**.
 2. whether the recording **contained any abnormal events (e.g., seizures)**.
 3. the **type of seizure** if present (complex partial, simple partial, absence, myoclonic, generalized tonic-clonic, other/none).
- As these tasks are a labeled subset of the SCORE specifications, we are using our system's performance on them as a **proxy for evaluating automatic SCORE extraction**.

TUH EEG Data

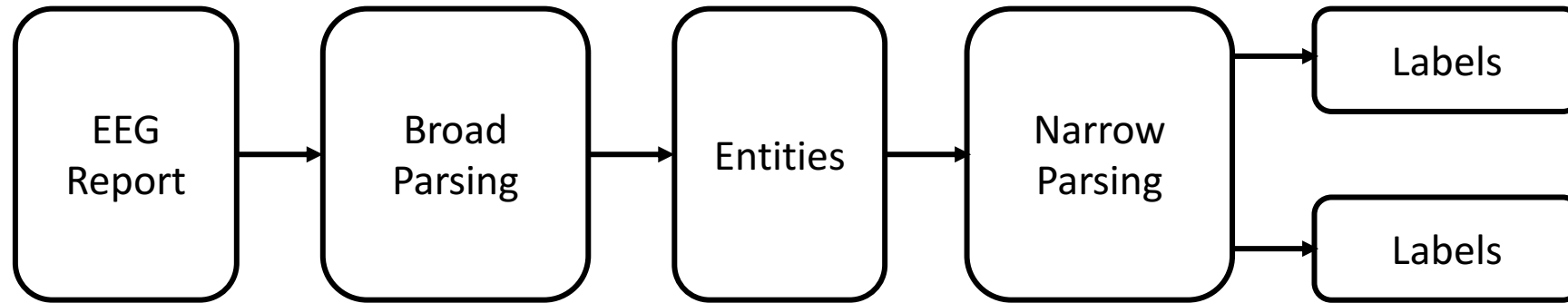
Dataset	Number of Samples
Epilepsy	561
Seizure	1105
Abnormal	2993

Related Work & Challenges

- Named Entity Recognition (NER) is a related classification task in the clinical domain where there is ample data availability.
 - Extraction of relevant “entities” (i.e. medication names, medical procedures, labs) from unstructured text.
 - Recent successful work on NER has been with fine-tuning Transformer deep neural networks such as BERT^[4].
- Challenges
 - Lack of training data (out of 16,000 TUH EEG records, certain classes have < 10 labels)
 - Varied clinician practices in describing clinical events or terms
 - “Record is NORMAL.”;
“No epileptiform features were seen.”
 - “Record is ABNORMAL.”
“This EEG is remarkable for...”

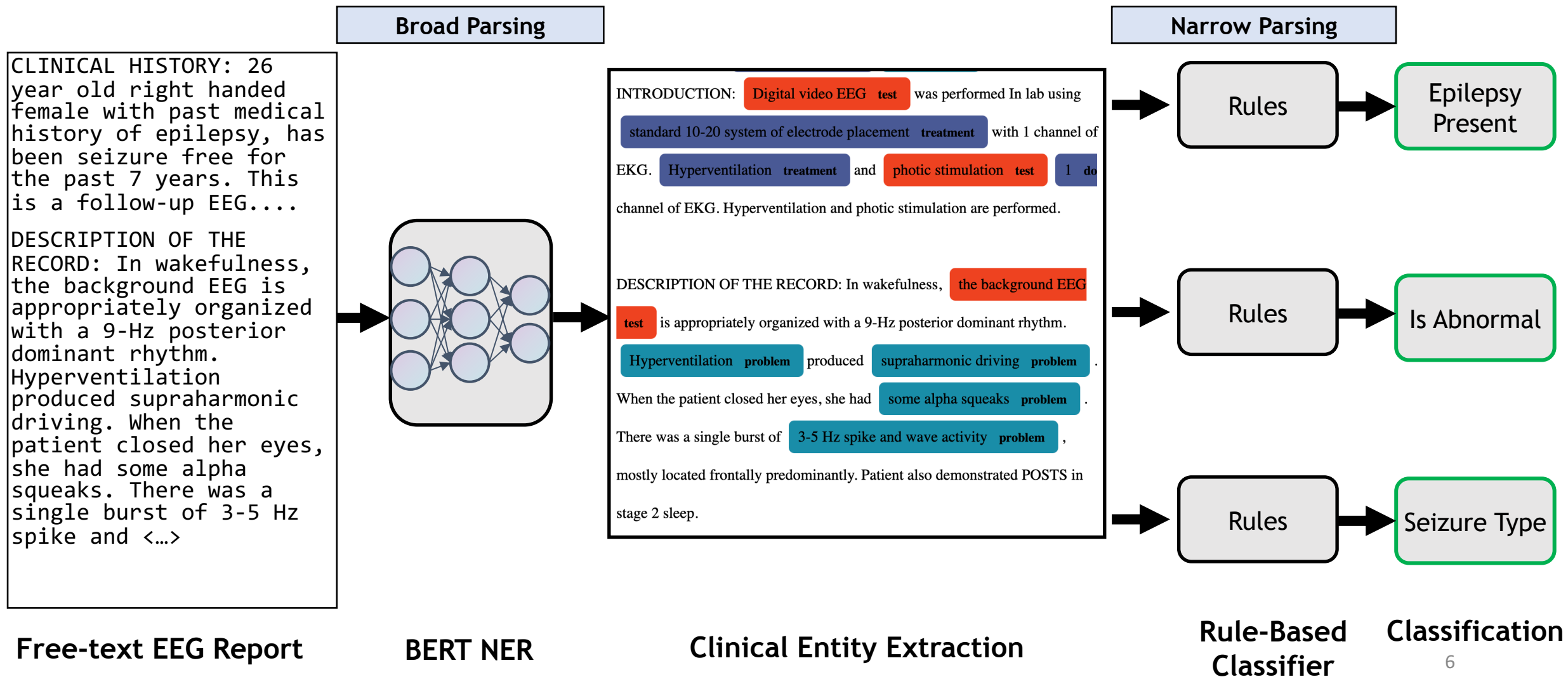
Dataset	Class	Train Support	Test Support
Seizure	Absence	10	6
	Complex Partial	45	13
	Myoclonic	1	12
	Simple Partial	2	0
	Tonic-Clonic	12	4
	None	913	97
Epilepsy	Epilepsy	428	428
	No Epilepsy	133	133
Abnormal	Normal	1371	150
	Abnormal	1346	126

Overview

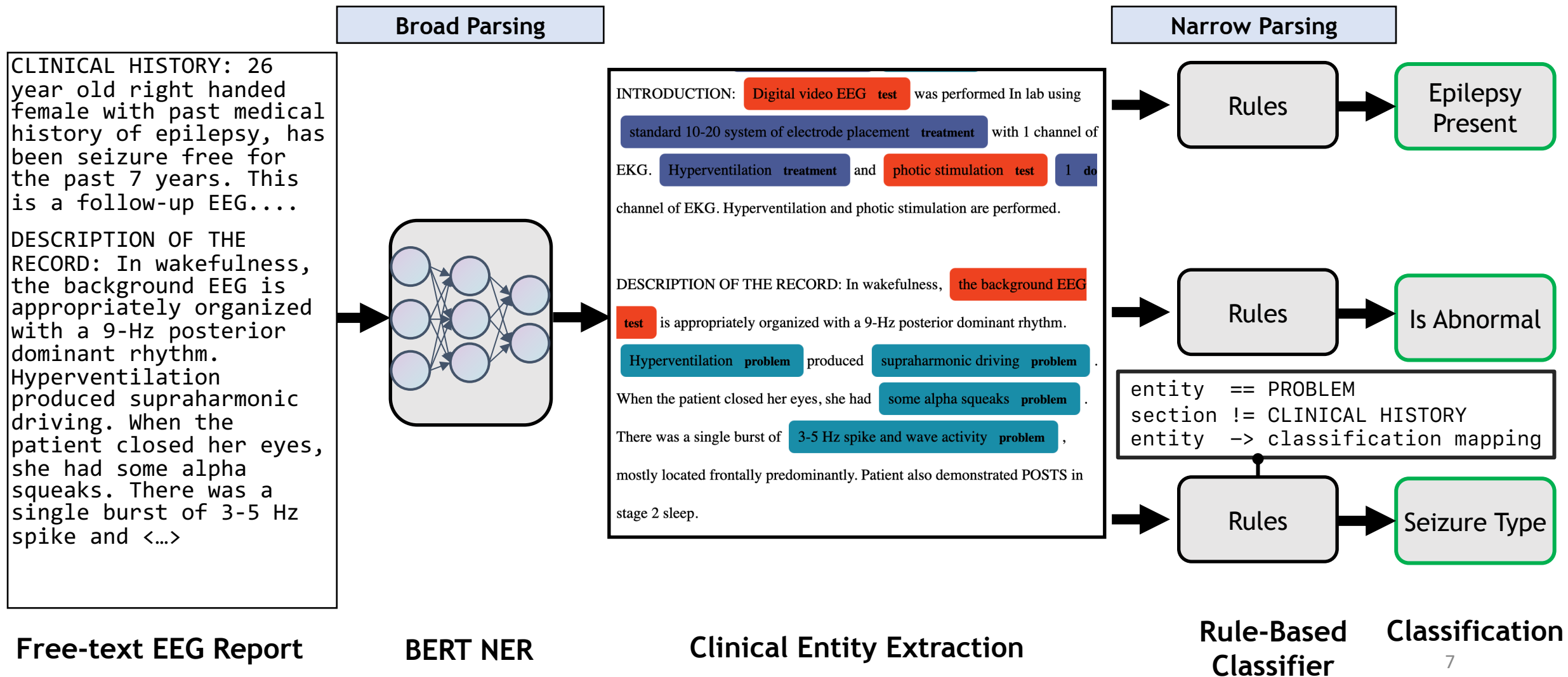


- Given unstructured document, the pipeline consists of two steps:
 - **“Broad” Feature Extraction:**
 - extraction of clinically relevant entities (i.e. medication, medical problems, treatments, labs).
 - via BERT-based deep neural network models finetuned on other available clinical NLP corpora.
 - **“Narrow” Classification:**
 - utilization of hand-crafted rules built on general entities to perform classification.
 - rules consist of medical record section header extraction, regular expressions, entity matching.

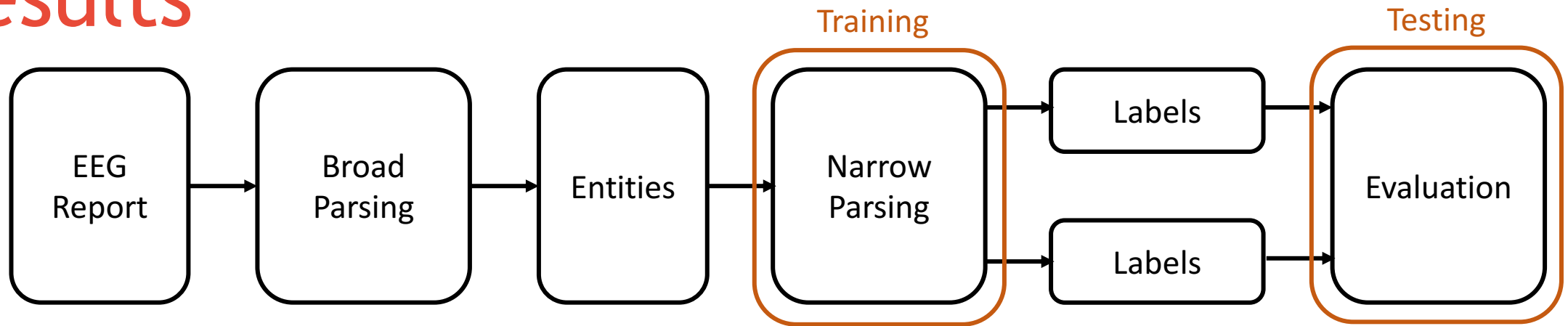
Methodology



Methodology



Results



Dataset

Dataset	Class	Train Support	Test Support
Seizure	Absence	10	6
	Complex Partial	45	13
	Myoclonic	1	12
	Simple Partial	2	0
	Tonic-Clonic	12	4
	None	913	97
Epilepsy	Epilepsy	428	428
	No Epilepsy	133	133
Abnormal	Normal	1371	150
	Abnormal	1346	126

System Performance

Task	Num Records	Weighted F1
Epilepsy Evaluation	561	0.82
Normal vs. Abnormal Classification	2727	0.97
Seizure Type Classification	171	0.92

Impacts & Future Work

Impacts

- Facilitate better indexing, searching, and organization of existing EEG reports.
- Automatic classification of free-text records for research purposes.

Future Goals

- Integration of clinical domain knowledge.
- Establishing ground truth standards.
- Robust verification of system.

References

- [1] S. Beniczky, H. Aurlien, J. C. Brøgger, A. Fuglsang-Frederiksen, A. Martins-da Silva, E. Trinka, G. Visser, G. Rubboli, H. Hjalgrim, H. Stefan et al., “Standardized computer-based organized reporting of eeg: Score,” *Epilepsia*, vol. 54, no. 6, pp. 1112- 1124, 2013.
- [2] A. Stubbs, M. Filannino, E. Soysal, S. Henry, and Ö. Uzuner, “Cohort selection for clinical trials: n2c2 2018 shared task track 1,” *Journal of the American Medical Informatics Association*, vol. 26, no. 11, pp. 1163-1171, 2019.
- [3] S. Rawal, A. Prakash, S. Adhya, S. Kulkarni, S. Anwar, C. Baral, and M. Devarakonda, “Semi-automated clinical lexicon induction and its use in cohort selection from clinical notes,” 2020 IEEE International Conference on Healthcare Informatics (ICHI). IEEE, 2020, pp. 1-2.
- [4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.
- [5] A. Harati, S. Lopez, I. Obeid, J. Picone, M. Jacobson, and S. Tobochnik, “The tuh eeg corpus: A big data resource for automated eeg interpretation,” 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). IEEE, 2014, pp. 1-5.