EMERGENCY MEDICAL SERVICE EMR-DRIVEN CONCEPT EXTRACTION FROM NARRATIVE TEXT

by

Susanna Serene George

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Electrical and Computer Engineering

Indianapolis, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Brian King, Co-Chair

Department of Electrical and Computer Engineering

Dr. Xiao Luo, Co-Chair

Department of Computer and Information Technology

Dr. Catharine Wyss

Department of Medical and Molecular Genetics

Approved by:

Dr. Brian King

Dedicated to my family who encouraged and supported me in all my work.

ACKNOWLEDGMENTS

As I begin, may all glory, honor and praise be to God for giving me an opportunity to learn, good health in the midst of a pandemic, and most importantly wisdom and understanding to do this work. I would like to express my sincere gratitude to Dr. Luo for her guidance and support in completing my research, and for her many and substantial contributions to this work. I also want to thank Dr. King and Dr. Wyss for offering me help based on their experience. And last but not the least, I am extremely grateful for my kind, amiable and supportive academic advisor, Sherrie Tucker for patiently answering all my queries, encouraging me and keeping me on track.

PREFACE

The basis of this research work started from a thought on how to use technology to help the first responders in the work they do. Done in collaboration with Dr.Zhan Zhang and Ricky Joseph Harris Jr. from Pace University New York, this work intends to provide relevant information regarding a patient to the hospital database once the EMS staff arrives at the scene and describes the situation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

advmod	adverbial modifier
amod	adjectival modifier
case	case marking
cc	coordinating conjunction
compound	compound
conj	conjunct
csubj	clausal subject
det	determiner
mark	marker
neg	negation modifier
nmod	nominal modifier
nsubj	nominal subject
nummod	numeric modifier
obj	object
obl	oblique nominal
punct	punctuation
CC	coordinating conjunction
DT	determiner
IN	preposition/subordinating conjunction
JJ	This NLTK POS Tag is an adjective (large)
NN	noun, singular (cat, tree)
RB	adverb (occasionally, swiftly)
RP	particle (about)
VB	verb (ask)
WP	wh- pronoun (who)

ABSTRACT

Being in the midst of a pandemic with patients having minor symptoms that quickly become fatal to patients with situations like a stemi heart attack, a fatal accident injury, and so on, the importance of medical research to improve speed and efficiency in patient care, has increased. As researchers in the computer domain work hard to use automation in technology in assisting the first responders in the work they do, decreasing the cognitive load on the field crew, time taken for documentation of each patient case and improving accuracy in details of a report has been a priority.

This paper presents an information extraction algorithm that custom engineers certain existing extraction techniques that work on the principles of natural language processing like metamap along with syntactic dependency parser like spacy for analyzing the sentence structure and regular expressions to recurring patterns, to retrieve patient-specific information from medical narratives. These concept value pairs automatically populates the fields of an EMR form which could be reviewed and modified manually if needed. This report can then be reused for various medical and billing purposes related to the patient.

1. INTRODUCTION

One of the highest priorities in the midst of a pandemic like the current COVID19 situation is the knowledge and expertise of hospital staff and EMS personnel. Emergency Medical Services(EMS) provides emergency medical care to patients more than 2,000 times per day[1]. With patients having minor injuries or symptoms, or patients that are old or having serious health conditions causing them to be either slow in speech or unable to speak, the EMS department is called upon by many different people with varying health concerns. In situations when a person has a stemi heart attack caused by a blocked coronary artery, as per the American Heart Association's guidelines, the blocked artery must be reopened within 90 minutes also known as the door to balloon time[2]. In such scenarios, speed and efficiency in thought and action is crucial to save the patient's life.

Research work in the field of emergency medical care and automation with the help of technology has improved patient care immensely by saving time during treatment and improving accuracy in visualizations, measurements and documentation or reporting. Since an EMS staff would have to deal with multiple cases of varying priorities throughout the day, it is essential that each patient case is documented to retain crucial information regarding the patient like vitals, cause of injury or illness, description of visuals from an accident site, state of patient before and after treatment provided by EMS and so on. With focus on providing the best treatment possible to the patient in a short time, it is often quite stressful and time-consuming for an EMS staff to prepare reports in parallel.

This research study emphasizes on reducing the cognitive overload for the emergency medical service personnel approaching the patient at a very critical time of their life. The combination of wearable devices that capture audio and images along with the work done as part of this paper will help generate a report from the conversation of an EMS staff at an incident site. With an EMS staff wearing a smart glass that provides an option to take pictures at the scene, scan medicine bar codes and record voice conversation, the description of the patient details like vitals, location of pain or injury, medications taken, past medical history, etc can be captured as an audio file which is processed using speech-to-text recognition software. The resulting output of a medical narrative is processed using the methodology described in this paper to produce the relevant concept names and their values specified in text. The concept-value pair is automatically populated to the EMR form and can be viewed by the staff. If any details are missing or incorrect, they have an option to add or change the values.

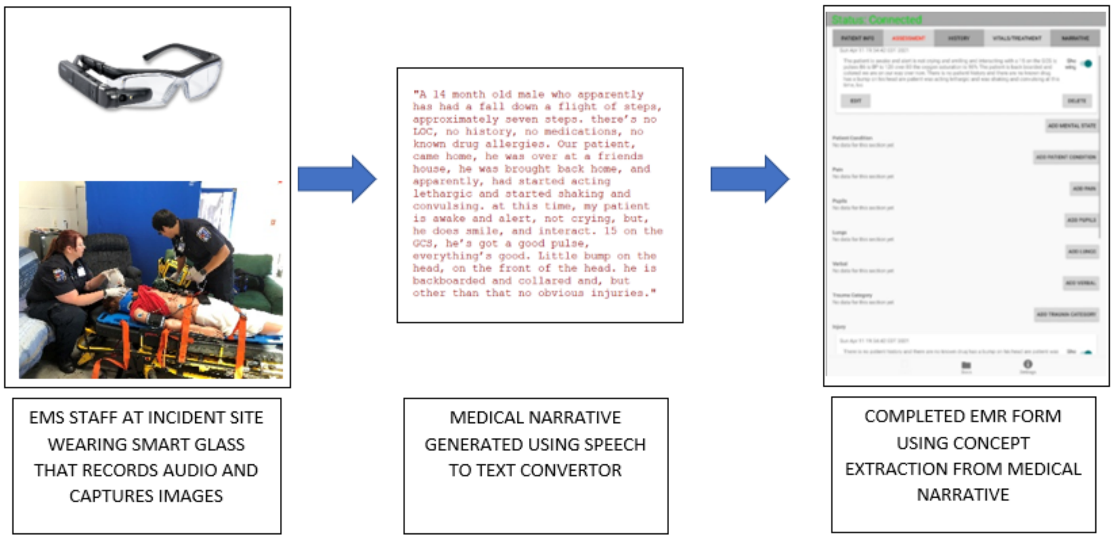
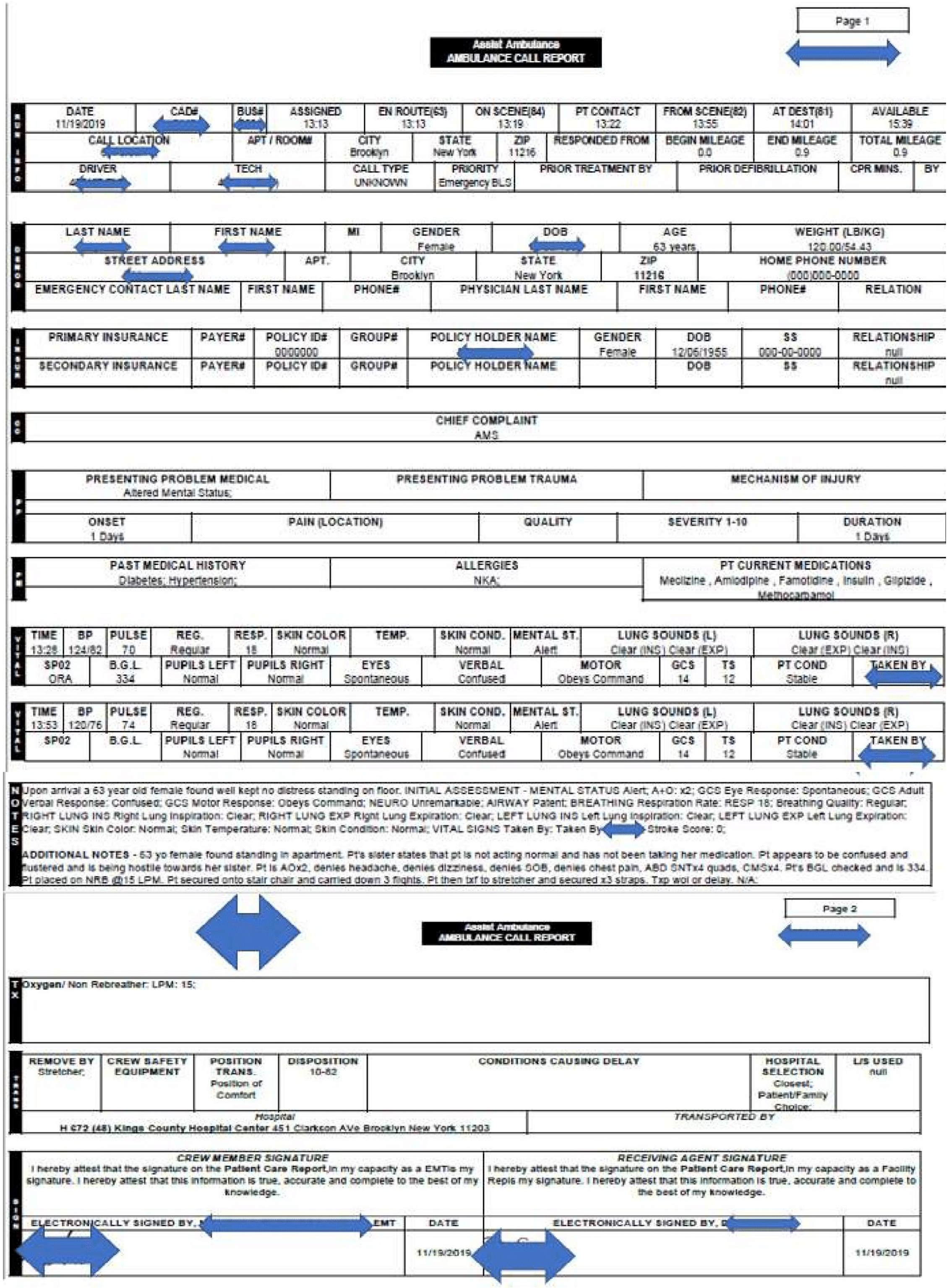


Figure 1.1. System Architecture[3][4]

Figure 1.2 shows a sample ambulance report. This report can then be provided to the ER nurse taking care of the patient or the doctor in advance through cloud based technology which would help the medical team at the hospital to make necessary arrangements to treat the patient as soon as possible. This report can also be later used for billing purposes, hence reducing the effort of the field crew in typing multiple reports and removing the need of recalling from memory the details of the case.



Page 1

Assist Ambulance
AMBULANCE CALL REPORT

DATE	CAD#	BUS#	ASSIGNED	EN ROUTE(63)	ON SCENE(84)	PT CONTACT	FROM SCENE(82)	AT DEST(81)	AVAILABLE
11/19/2019	[redacted]	[redacted]	13:13	13:13	13:19	13:22	13:55	14:01	15:39

CALL LOCATION	APT / ROOM#	CITY	STATE	ZIP	RESPONDED FROM	BEGIN MILEAGE	END MILEAGE	TOTAL MILEAGE
[redacted]		Brooklyn	New York	11216		0.0	0.9	0.9

DRIVER	TECH	CALL TYPE	PRIORITY	PRIOR TREATMENT BY	PRIOR DEFIBRILLATION	CPR MINS.	BY
[redacted]	[redacted]	UNKNOWN	Emergency BLS				

LAST NAME	FIRST NAME	MI	GENDER	DOB	AGE	WEIGHT (LB/KG)
[redacted]	[redacted]		Female	[redacted]	63 years	120.00/54.43

STREET ADDRESS	APT.	CITY	STATE	ZIP	HOME PHONE NUMBER
[redacted]		Brooklyn	New York	11216	(000)000-0000

EMERGENCY CONTACT LAST NAME	FIRST NAME	PHONE#	PHYSICIAN LAST NAME	FIRST NAME	PHONE#	RELATION

PRIMARY INSURANCE	PAYER#	POLICY ID#	GROUP#	POLICY HOLDER NAME	GENDER	DOB	SS	RELATIONSHIP
		0000000		[redacted]	Female	12/06/1955	000-00-0000	null
SECONDARY INSURANCE	PAYER#	POLICY ID#	GROUP#	POLICY HOLDER NAME		DOB	SS	RELATIONSHIP
								null

CHIEF COMPLAINT
AMS

PRESENTING PROBLEM MEDICAL	PRESENTING PROBLEM TRAUMA	MECHANISM OF INJURY
Altered Mental Status;		

ONSET	PAIN (LOCATION)	QUALITY	SEVERITY 1-10	DURATION
1 Days				1 Days

PAST MEDICAL HISTORY	ALLERGIES	PT CURRENT MEDICATIONS
Diabetes; Hypertension;	NKA;	Meclizine , Amlodipine , Famotidine , Insulin , Glipizide , Methocarbamol

TIME	BP	PULSE	REG.	RESP.	SKIN COLOR	TEMP.	SKIN COND.	MENTAL ST.	LUNG SOUNDS (L)	LUNG SOUNDS (R)
13:28	124/82	70	Regular	18	Normal		Normal	Alert	Clear (INS) Clear (EXP)	Clear (EXP) Clear (INS)

SP02	B.G.L.	PUPILS LEFT	PUPILS RIGHT	EYES	VERBAL	MOTOR	GCS	TS	PT COND	TAKEN BY
ORA	334	Normal	Normal	Spontaneous	Confused	Obeys Command	14	12	Stable	[redacted]

TIME	BP	PULSE	REG.	RESP.	SKIN COLOR	TEMP.	SKIN COND.	MENTAL ST.	LUNG SOUNDS (L)	LUNG SOUNDS (R)
13:53	120/76	74	Regular	18	Normal		Normal	Alert	Clear (INS) Clear (EXP)	Clear (INS) Clear (EXP)

SP02	B.G.L.	PUPILS LEFT	PUPILS RIGHT	EYES	VERBAL	MOTOR	GCS	TS	PT COND	TAKEN BY
		Normal	Normal	Spontaneous	Confused	Obeys Command	14	12	Stable	[redacted]

NOTES

Upon arrival a 63 year old female found well kept no distress standing on floor. INITIAL ASSESSMENT - MENTAL STATUS Alert; A+O: x2; GCS Eye Response: Spontaneous; GCS Adult Verbal Response: Confused; GCS Motor Response: Obeys Command; NEURO Unremarkable; AIRWAY Patent; BREATHING Respiration Rate: RESP 18; Breathing Quality: Regular; RIGHT LUNG INS Right Lung Inspiration: Clear; RIGHT LUNG EXP Right Lung Expiration: Clear; LEFT LUNG INS Left Lung Inspiration: Clear; LEFT LUNG EXP Left Lung Expiration: Clear; SKIN Skin Color: Normal; Skin Temperature: Normal; Skin Condition: Normal; VITAL SIGNS Taken By: Taken By [redacted] Stroke Score: 0;

ADDITIONAL NOTES - 63 yo female found standing in apartment. Pt's sister states that pt is not acting normal and has not been taking her medication. Pt appears to be confused and flustered and is being hostile towards her sister. Pt is AOx2, denies headache, denies dizziness, denies SOB, denies chest pain, ABD SNTx4 quads, CMSx4. Pt's BGL checked and is 334. Pt placed on NRB @15 LPM. Pt secured onto stair chair and carried down 3 flights. Pt then txf to stretcher and secured x3 straps. Txp wol or delay. N/A:

[redacted]

Page 2

Assist Ambulance
AMBULANCE CALL REPORT

TX

Oxygen/ Non Rebreather: LPM: 15;

REMOVE BY	CREW SAFETY EQUIPMENT	POSITION TRANS.	DISPOSITION	CONDITIONS CAUSING DELAY	HOSPITAL SELECTION	L/S USED
Stretcher;		Position of Comfort	10-82		Closest; Patient/Family Choice;	null

Hospital	*TRANSPORTED BY*
H 672 (48) Kings County Hospital Center 451 Clarkson AVe Brooklyn New York 11203	

CREW MEMBER SIGNATURE	*RECEIVING AGENT SIGNATURE*
I hereby attest that the signature on the **Patient Care Report**,in my capacity as a EMTis my signature. I hereby attest that this information is true, accurate and complete to the best of my knowledge.	I hereby attest that the signature on the **Patient Care Report**,in my capacity as a Facility Repis my signature. I hereby attest that this information is true, accurate and complete to the best of my knowledge.

ELECTRONICALLY SIGNED BY,	EMT	DATE	ELECTRONICALLY SIGNED BY,	DATE
[redacted]		11/19/2019	[redacted]	11/19/2019

Figure 1.2. Sample Interface of Ambulance Call Report

2. RELATED WORK

Concept extraction is the process of identifying certain concepts in a specific domain and extracting information relevant to each concept as words, phrases or short sentences from a large text data. In other words, it is a combination of concept mention detection and concept encoding. Concept mention detection generally adopts the named entity recognition(NER) technology in the general domain, which focuses on detecting concept mentions in the text and concept encoding aims to map the mentions to concepts in standard terminologies[5].

Information extraction from resume uses a cascade model. In the first pass, a resume is segmented into a series of blocks attached with labels indicating the type of content followed by the second pass during which information such as Name and Address, are identified in certain blocks and then the most appropriate model is selected through experiments for each IE task in different passes[6]. Mermelstein[7] et al. uses an integration of hidden Markov models and text classifiers to extract information from business cards and email messages requesting change of address. DBpedia uses an infobox extraction algorithm that detects certain specific templates and it's structure using pattern matching techniques[8]. It selects significant templates, which are then parsed and transformed to Resource Description Framework(RDF) triples[8]. PROSPERA harvests high-quality knowledge using ngram-itemsets for richer patterns and Maximum Satisfiability based constraint reasoning on both the quality of patterns and the validity of fact candidates[9].

Natural language processing(NLP) systems have been used to generate concepts as structured information from unstructured free text. NLP techniques model the way a user requests information to how computer or software understands it. In Social Network monitoring systems that track user interactions and communications, information extraction plays a key role. Charu et al. uses data mining techniques that involve association, correlation, classification, cluster analysis and linkage based cross learning to monitor and enhance user experience in Social Media[10]. In healthcare, most of the existing information extraction (IE) methods targeting medical reports are either dictionary-based or rule-based and require the targeted types of relations and the synonymous relations to be pre-defined. These are extremely time consuming which result in a low recall as the manual rules or patterns are required to be pre-defined by a domain expert[11].

There is an emerging trend of hybrid systems that perform well in clinical information retrieval, with rule-based NLP being the most commonly used approach followed by supervised machine learning approaches with one disadvantage being the need of large amounts of annotated data[11].

The generic concept extraction techniques cannot be directly applied to medical text as it often contains ambiguity, polysemy, and variation of word orders[12]. Hence, several existing research are customized to narrow down focus to information extraction methods from the medical text. An SVM classifier uses lexical, medical, and structural features to extract patient's clinical problems from electronic medical records[12]. Since the EMS data contains terminology unique to the EMS department, the aforementioned research can not be directly applied for EMS concept extraction. Also, these approaches often rely on classifiers with only a few class of concepts with a reasonable amount of annotated data for each class. However, the target set of EMS concepts can not be unambiguously categorized in a few classes, and the annotation is very expensive[12].

Documentation of each patient case by the field crew has significantly improved once the paper report forms have been converted to electronic reports that are guided by some of the currently available reporting tools. Although ImageTrend1[13] is an increasingly popular tool for documentation, tracking and visualization of EMS information and Emergency Department Information Exchange (EDIE)[14] links all hospital emergency departments by facilitating real-time communication and collaboration, both ImageTrend and EDIE, require manual input in the initial phase of data collection which is tedious and prone to errors[15].Clinical Language Annotation, Modeling, and Processing Toolkit(CLAMP) is a comprehensive clinical Natural Language Processing (NLP) software that enables recognition and automatic encoding of clinical information in narratives[16]. MetaMap[17], cTAKES[18] and CLAMP use the Unified Medical Language System (UMLS) to extract medical concepts. While these systems have been in use for a long time now, it has not been completely successful in obtaining the contexts in finer granularity. For example, MetaMap has Concept Unique Identifiers (CUI) and semantic type lists which signify whether a clinical concept is 'Disease' or 'Medication' but it does not differentiate if the medication is given to the patient by another person like the EMS staff or whether it is the medication that is taken by the patient for a pre-existing condition. Generating Summary Reports Automatically for Cognitive Assistance in Emergency Response(GRACE) does not require any such effort, as summary reports are automatically generated from the audio data captured on-scene and processing it using NLP[15]. This research paper intends to provide efficient and accurate concept extraction from the ems medical narratives.

As part of text extraction, this research paper also compares some of the most common multi-label classification models to classify medical narratives based on the mechanism of injury described in the text. Naive Bayes classifiers that work on the principle of Bayes theorem assumes that each pair of features are independent of each other and generates a hypothesis based on some evidence in hand. Although the assumption of independence is not correct in real-life scenarios, naive bayes classifiers do produce significant results compared to some other classifiers[19]. With a larger vocabulary, the size of input and processing time increases which leads us to a model like Word2Vec which provides a fixed size representation of words that is much smaller in size compared to the size of the vocabulary. Word2Vec model[20] combined a recurrent neural network classifier such as BiLSTM produces better results in a shorter time. BiLSTM models are used for purposes like classification, sequence tagging[21] and more. Finally, the state-of-the-art language model BERT used for various purposes like classification, Q&A, sentiment analysis, prediction and more was considered. One of the most common facts known to data scientists or researchers in the field of machine learning is that performance mostly improves with size of dataset used for training and that is one of the salient features of BERT along with many others with the most important one being it's bidirectional nature[22]. Since BERT is trained on simple plain text from various domains, it's efficiency in classifying emergency medical care focused narratives was comparatively poor although the data was duplicated to provide a balanced input. The lighter version of BERT, DistilBERT was also considered for classification but the results were not as expected and a possible explanation concluded was the difference in domain used for training the easily available pre-trained embeddings with respect to the EMS domain used for fine-tuning the embeddings.

3. METHODOLOGY

3.1 Information Extraction from Medical Narrative

As the first responder arrives at the incident scene and reports the medical information regarding the patient, the information is captured using the wearable device and then converted to text which is then processed by the system with architecture shown in Figure 3.1, to retrieve only the relevant details.

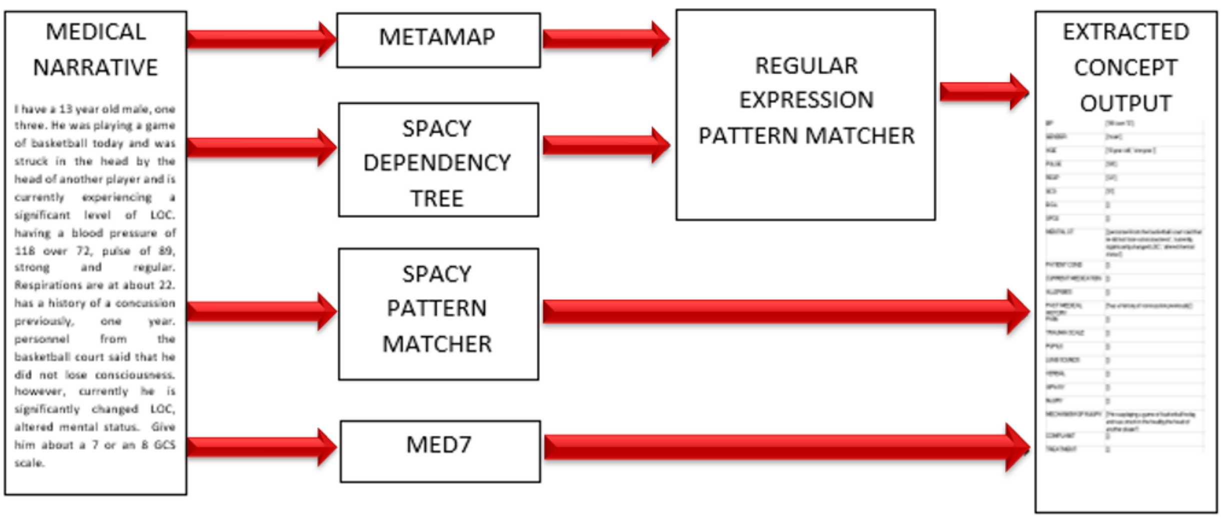


Figure 3.1. Architecture Framework

3.1.1 Medical Narrative

The current standard for documentation of care by the emergency medical services (EMS) personnel is the paper-based patient care report (PCR) or the electronic PCR (ePCR) which is often incomplete at the time of transfer of care to the emergency department (ED) and the same information has to be entered in several places to satisfy the requirements of all entities involved in EMS care and public health [23] . Provided with a detailed narrative that may or may not follow the DACHARTE or CHART documentation format which describes the information from the ECC on dispatch, scene on arrival, patient complaint, history of present condition, patient assessment, treatment provided, transport mechanism and exceptions if any [24] , this paper intends to retrieve relevant information to generate an electronic PCR using natural language processing (NLP) techniques.

3.1.2 METAMAP

MetaMap, an application which uses knowledge-intensive approach, natural language processing and computational-linguistic techniques, is employed to map biomedical text to the UMLS Metathesaurus and identify Metathesaurus concepts referred to in the provided narrative report from the EMS personnel [25] . A brief look into the algorithm behind the working of the MetaMap leads us to 6 major steps with added features that can be configured during execution.

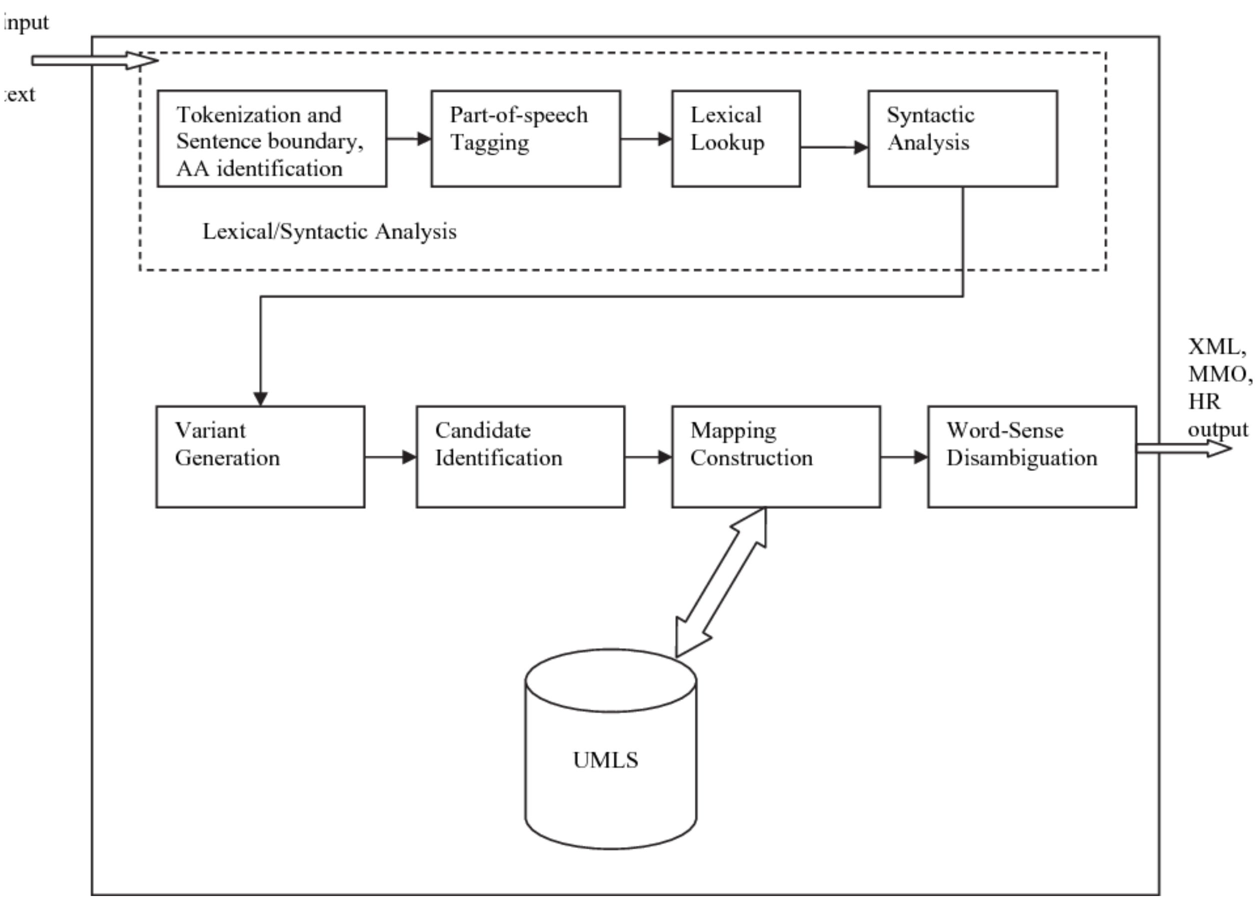


Figure 3.2. Metamap Algorithm [26]

Tokenization

A tokenizer uses a set of rules to understand the language of the text, detect sentence boundaries, identify tokens in text, characters of punctuation and acronyms or abbreviations(AA)[27].

Parsing

The text is parsed into phrases which makes the mapping effort more controllable. Parsing is performed using the SPECIALIST minimal commitment parser [28] which produces a shallow syntactic analysis of the text. The parser uses the Xerox part-of-speech tagger [29] which assigns syntactic tags (e.g., noun, verb) to words not having a unique tag in the SPECIALIST lexicon.[25]

Variant Generation

```
drug [noun] variants (n=8):
drug{[noun], 0=[]}
medicamentous{[adj], 2="s"}
pharmaceutical{[adj], 2="s"}
drugs{[noun], 4="ss"}
medicament{[noun], 5="ds"}
medicaments{[noun], 5="ds"}
medicamental{[adj], 8="dds"}
medicamented{[adj], 8="dds"}
```

Figure 3.3. Variants of *drug*

For each phrase, variants are generated using the knowledge in the SPECIALIST lexicon and a supplementary database of synonyms. A variant consists of a phrase word (called a generator) together with all its acronyms, abbreviations, synonyms, derivational variants, meaningful combinations of these, and finally inflectional and spelling variants[25]. For example, Figure 3.3 shows the variants generated for the word 'drug' in the phrase 'no known drug allergies'.

Candidate Retrieval

Intermediate results consisting of Metathesaurus strings, called candidates, matching some phrase text are computed[26] and retrieved[25].

```
1000   No known drug allergies (No known drug allergy) [Finding]
 923   NO KNOWN ALLERGIES (No known allergy) [Finding]
 906   NO DRUG ALLERGY (no drug allergy) [Finding]
 861 N Drug Allergies (Drug Allergy) [Pathologic Function]
 858 N Known allergy [Finding]
 827 E Drug allergy (Drug Allergy Allergen) [Immunologic Factor]
 812   Drug (Pharmaceutical Preparations) [Pharmacologic Substance]
 812 N Allergies (Hypersensitivity) [Pathologic Function]
 812   DRUG (Pharmacologic Substance) [Pharmacologic Substance]
 779 E Allergy (Allergy Specialty) [Biomedical Occupation or Discipline]
 779 E ALLERGY (Response to antigens) [Physiologic Function]
 779 E Allergy (Allergic Reaction) [Pathologic Function]
 779 E Allergy (Allergic disposition) [Finding]
 756 E Pharmaceutical (Pharmacy (field)) [Biomedical Occupation or Discipline]
 756 E pharmaceutical (Procedures involving the use of pharmaceuticals) [Therapeutic or Preventive Procedure]
 729 E Allergic [Functional Concept]
 729 E Drugs (Drugs - dental services) [Therapeutic or Preventive Procedure]
 719 E Medicaments (Medicament) [Pharmacologic Substance]
 645   NO (nitric oxide) [Organic Chemical,Pharmacologic Substance]
 645   Known [Qualitative Concept]
 645   No (no) [Finding]
 645   No (Negation) [Functional Concept]
 645   No (No - yes/no indicator) [Idea or Concept]
 645   NO (Nitric Oxide Measurement) [Laboratory Procedure]
```

Figure 3.4. Metathesaurus candidates for *no known drug allergies*

Candidate Evaluation

Each Metathesaurus candidate is evaluated against the input text by first computing a mapping from the phrase words to the candidate's words and then calculating the strength of the mapping using a linguistically principled evaluation function consisting of a weighted average of four metrics: centrality (involvement of the head), variation (an average of inverse distance scores), coverage and cohesiveness[25]. The latter two components measure how much of a candidate matches the text and in how many pieces[25]. The candidates are then ordered according to mapping strength.[25]

Mapping Construction

Complete mappings are constructed by combining candidates involved in disjoint parts of the phrase, and the strength of the complete mappings is computed just as for candidate mappings. The highest scoring complete mappings represent MetaMap's best interpretation of the original phrase[25]. In our study we use candidates that have a score greater than or equal to 800 as our threshold limit.

Semantic Type

The output from Metamap is then filtered based on required semantic type[30] from a set of about 127 different semantic types with respect to the context of the data that is to be retrieved.

```
aapp|T116|Amino Acid, Peptide, or Protein
acab|T020|Acquired Abnormality
acty|T052|Activity
aggp|T100|Age Group
amas|T087|Amino Acid Sequence
amph|T011|Amphibian
anab|T190|Anatomical Abnormality
anim|T008|Animal
anst|T017|Anatomical Structure
antb|T195|Antibiotic
arch|T194|Archaeon
bacs|T123|Biologically Active Substance
bact|T007|Bacterium
bdsu|T031|Body Substance
bdsy|T022|Body System
bhvr|T053|Behavior
biof|T038|Biologic Function
bird|T012|Bird
blor|T029|Body Location or Region
bmod|T091|Biomedical Occupation or Discipline
bodm|T122|Biomedical or Dental Material
bpoc|T023|Body Part, Organ, or Organ Component
bsoj|T030|Body Space or Junction
celc|T026|Cell Component
celf|T043|Cell Function
cell|T025|Cell
cgab|T019|Congenital Abnormality
chem|T103|Chemical
chvf|T120|Chemical Viewed Functionally
chvs|T104|Chemical Viewed Structurally
clas|T185|Classification
clna|T201|Clinical Attribute
clnd|T200|Clinical Drug
cnce|T077|Conceptual Entity
comd|T049|Cell or Molecular Dysfunction
crbs|T088|Carbohydrate Sequence
diap|T060|Diagnostic Procedure
dora|T056|Daily or Recreational Activity
drdd|T203|Drug Delivery Device
dsyn|T047|Disease or Syndrome
edac|T065|Educational Activity
eehu|T069|Environmental Effect of Humans
elii|T196|Element, Ion, or Isotope
emod|T050|Experimental Model of Disease
emst|T018|Embryonic Structure
enty|T071|Entity
enzy|T126|Enzyme
euka|T204|Eukaryote
evnt|T051|Event
famg|T099|Family Group
ffas|T021|Fully Formed Anatomical Structure
fish|T013|Fish
fndg|T033|Finding
fngs|T004|Fungus
food|T168|Food
ftcn|T169|Functional Concept
genf|T045|Genetic Function
geoa|T083|Geographic Area
gngm|T028|Gene or Genome
gora|T064|Governmental or Regulatory Activity
grpa|T102|Group Attribute
grup|T096|Group
hcpp|T068|Human-caused Phenomenon or Process
hcro|T093|Health Care Related Organization
hlca|T058|Health Care Activity
hops|T131|Hazardous or Poisonous Substance
horm|T125|Hormone
humn|T016|Human
idcn|T078|Idea or Concept
imft|T129|Immunologic Factor
inbe|T055|Individual Behavior
inch|T197|Inorganic Chemical
inpo|T037|Injury or Poisoning
inpr|T170|Intellectual Product
```

Figure 3.5. Metamap Semantic Types

3.1.3 Syntactic Dependency Tree

The Dependency parser, spacy is a free, open-source library designed specifically for production use to build information extraction or natural language understanding systems, or to pre-process text for deep learning[31]. ScispaCy is a Python package containing spaCy models for processing biomedical, scientific or clinical text efficiently built by retraining spaCy3 models for POS tagging, dependency parsing, and NER using datasets relevant to biomedical text, and enhancing the tokenization module with additional rules[32]. The spacy model, 'en_ner_bc5cdr_md' trained on the BC5CDR corpus mainly recognizes 2 main additional entity types compared to the basic spacy models namely disease and chemical entity types[33]. The BC5CDR corpus consists of 1500 PubMed articles with 4409 annotated chemicals, 5818 diseases and 3116 chemical-disease interactions, developed for use in the BioCreative V challenge tasks of disease named entity recognition(DNER) and chemical-induced disease(CID) relation extraction and also serves as a valuable resource for the text-mining community[34].

Using the syntactic dependency parser, token tags that contain information about the word types like the parts of speech and token dependencies that describe how the words are related to each other in a sentence, are obtained in the form of a dependency tree. A dependency tree is a directed graph that represent the grammatical structure of a sentence or phrase which delineates the dependency between a word (such as a verb) and the phrases it builds upon (such as the subject and object phrases of that verb)[35][36] A language model is loaded and then the sentence text or document is processed using the spacy language model which in the case of medical text used in our study is 'en_ner_bc5cdr_md' model.

During processing, spaCy first tokenizes the text into a Doc, i.e. segments it into words, punctuation and so on based on rules specific to each language[37]. The raw text is split on whitespace characters after which the tokenizer processes the text from left to right, checking whether each substring matches a tokenizer exception rule and if a prefix, suffix or infix can be split off, eventually letting spacy split complex, nested tokens like combinations of abbreviations and multiple punctuation marks[37].

After tokenization, spaCy parses and tags a given Doc with the help of a trained pipeline and its statistical models, to make predictions of which tag or label most likely applies to the word in the context given [37]. Part-of-speech(POS) tagging assigns word types to tokens, like verb or noun and dependency parsing assigns syntactic dependency labels, describing the relations between individual tokens, like subject or object [37].

TEXT	LEMMA	POS	TAG	DEP	SHAPE	ALPHA	STOP
A	a	DET	DT	det	X	TRUE	TRUE
14	14	NUM	CD	nummod	dd	FALSE	FALSE
month	month	NOUN	NN	npadvmod	xxxx	TRUE	FALSE
old	old	ADJ	JJ	amod	xxx	TRUE	FALSE
male	male	NOUN	NN	ROOT	xxxx	TRUE	FALSE
who	who	PRON	WP	nsubj	xxx	TRUE	TRUE
apparently	apparently	ADV	RB	advmod	xxxx	TRUE	FALSE
has	have	AUX	VBZ	aux	xxx	TRUE	TRUE
had	have	VERB	VBN	relcl	xxx	TRUE	TRUE
a	a	DET	DT	det	x	TRUE	TRUE
fall	fall	NOUN	NN	dobj	xxxx	TRUE	FALSE
down	down	ADP	IN	prep	xxxx	TRUE	TRUE
a	a	DET	DT	det	x	TRUE	TRUE
flight	flight	NOUN	NN	pobj	xxxx	TRUE	FALSE
of	of	ADP	IN	prep	xx	TRUE	TRUE
steps	step	NOUN	NNS	pobj	xxxx	TRUE	FALSE
,	,	PUNCT	,	punct	,	FALSE	FALSE
approximate	approximate	ADV	RB	advmod	xxxx	TRUE	FALSE
seven	seven	NUM	CD	nummod	xxxx	TRUE	FALSE
steps	step	NOUN	NNS	appos	xxxx	TRUE	FALSE
.	.	PUNCT	.	punct	.	FALSE	FALSE
There	there	PRON	EX	expl	Xxxxx	TRUE	TRUE
's	be	AUX	VBZ	ROOT	'x	FALSE	TRUE
no	no	DET	DT	det	xx	TRUE	TRUE
LOC	LOC	PROPN	NNP	attr	XXX	TRUE	FALSE
,	,	PUNCT	,	punct	,	FALSE	FALSE
no	no	DET	DT	det	xx	TRUE	TRUE
history	history	NOUN	NN	appos	xxxx	TRUE	FALSE
,	,	PUNCT	,	punct	,	FALSE	FALSE
no	no	DET	DT	det	xx	TRUE	TRUE
medications	medication	NOUN	NNS	appos	xxxx	TRUE	FALSE

Figure 3.6. Parsed and tagged text

Figure 3.6 provides a detailed view of the output provided after parsing and tagging the sentence *A 14 month old male who apparently has had a fall down a flight of steps,* *approximately seven steps. There's no LOC, no history, no medications* where TEXT represents the original word text, LEMMA gives the original base form of the word, POS stands for part-of-speech, TAG gives the detailed information of the part-of-speech, DEP is the syntactic dependency, i.e the relation between tokens, SHAPE describes the word shape – capitalization, punctuation and digits, ALPHA tells if the token is an alpha character or not and STOP provides information regarding whether the token is part of a stop list which is the most common words of the language. This information can then be used to identify contextual words describing certain entities based on dependencies.

Each word could either be a head/parent or a child to another word or both. The connection between words also known as the type of syntactic relation between words is represented by a single arc with a dependency label. Since the relations between words form a tree where each word has exactly one head, by iterating over each word of the sentence each of the dependencies or connections between the words of the sentence can be retrieved.

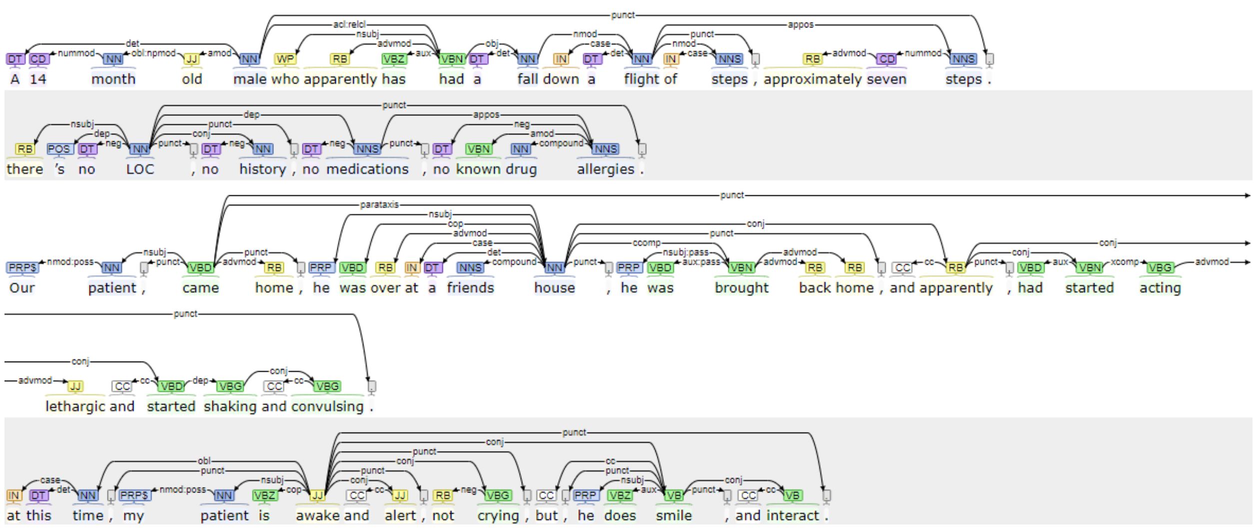


Figure 3.7. Dependency Tree[38]

Figure 3.7 shows the dependency tree for the paragraph '*A 14 month old male who apparently has had a fall down a flight of steps, approximately seven steps. There's no LOC, no history, no medications, no known drug allergies. Our patient, came home, he was over at a friend's house, he was brought back home, and apparently, had started acting lethargic and started shaking and convulsing. At this time, my patient is awake and alert, not crying,* *but, he does smile, and interact.*' where the colored box labels represent the part of speech tag like NN for noun, DT for determiner, CD for cardinal digit, JJ for adjective, WP for Wh-pronoun and so on[39] and the arc labels represent the dependency between the starting head word and the connected children. For example, in the phrase '*my patient is awake and alert, not crying, but, he does smile and interact*', the word 'awake' is the head/root word connected to the punctuations ',' and '.' with dependency label punctuation(punct) and words

- 'patient' with dependency label as nominal subject(nsubj) which implies it is the syntactic subject of the clause[40]
- 'alert', 'crying' and 'smile' with dependency label as conjunct(conj) which implies that they are connected by a coordinating conjunction

3.1.4 Regular Expression Pattern Matcher

Regular expressions also known as REs, or regexes, or regex patterns compiled into a series of bytecodes and then executed by a matching engine written in C is a highly specialized programming language embedded inside Python and made available through the re module to specify the rules for the set of possible strings to be matched which might contain English sentences, or e-mail addresses, or any other relevant phrases that occur in a predictable pattern[41].

Flag	Meaning
`ASCII`, `A`	Makes several escapes like `\w`, `\b`, `\s` and `\d` match only on ASCII characters with the respective property.
`DOTALL`, `S`	Make `.` match any character, including newlines.
`IGNORECASE`, `I`	Do case-insensitive matches.
`LOCALE`, `L`	Do a locale-aware match.
`MULTILINE`, `M`	Multi-line matching, affecting `^` and `$`.
`VERBOSE`, `X` (for 'extended')	Enable verbose REs, which can be organized more cleanly and understandably.

Figure 3.8. Regex Flags[41]

Most of the alphabets and digits characters and a few of the symbols will simply match themselves. For example, the regular expression *male* will match the string *male* exactly and if flag option 'IGNORECASE' is used, matched strings would be *male, MALE, Male, MAle, etc.*The exception to this rule are special natured metacharacters described in Figure 3.9.

Metacharacter	Usage
.	used to match any character except a newline character.
[]	used for specifying a character class which is a set of characters listed individually, or as a range of characters specified by two characters seperated by '-'. For example, [a-z] will match any of the characters from a to z.Metacharacters inside a character class is stripped of its special nature.
^	used at the beginning to complement a set. Example, [^0-9] will match any non-digit character.
$	used to match the end of a line defined as the end of the string, or any location followed by a newline character.
\	used to escape metacharacters and when followed by certain characters, it can also be used to signal a category or group of characters.
\w	equivalent to the class [a-zA-Z0-_] matches any alphanumeric character
\W	equivalent to the class [^a-zA-Z0-9_] matches any non-alphanumeric character
\d	equivalent to the class [0-9] matches any decimal digit
\D	equivalent to the class [^0-9] matches any non-digit character
\s	equivalent to the class [\t\n\r\f\v] matches any whitespace character
\S	equivalent to the class [^ \t\n\r\f\v] matches any non-whitespace character
*	used to match the previous character zero or more times rather
+	used to match the previous character one or more times
?	used to match the previous character zero or one time
{m,n}	used to match the previous character at least m times and at most n times
()	used to match a group of characters
|	used to match pattern on either side

Figure 3.9. Regex Metacharacters

The regular expression is compiled to obtain a pattern object that has various methods which can be used to identify if the pattern specified exists in the start of the string or any other part of the string, retrieve part of the matched string or the complete matched string, retrieve all occurrences of the matched string as a list or iterator, substitute the matched pattern with a different substring or part of the current string using groups. While trying to match patterns, the matching engine goes as far as it can at first, and on the off chance that no match is found it will then progressively back up and retry the rest of the RE again and again[41].

Method/Attribute	Purpose
`match()`	Determine if the RE matches at the beginning of the string.
`search()`	Scan through a string, looking for any location where this RE matches.
`findall()`	Find all substrings where the RE matches, and returns them as a list.
`finditer()`	Find all substrings where the RE matches, and returns them as an iterator.
`split()`	Split the string into a list, splitting it wherever the RE matches
`sub()`	Find all substrings where the RE matches, and replace them with a different string
`subn()`	Does the same thing as `sub()`, but returns the new string and the number of replacements

Figure 3.10. Regex Methods[41]

Similar methods are also available as module-level functions. Outside of loops, there's not much difference between module-level functions and invoking methods through pattern objects as compiled objects are stored in cache and can be reused by future calls [41].

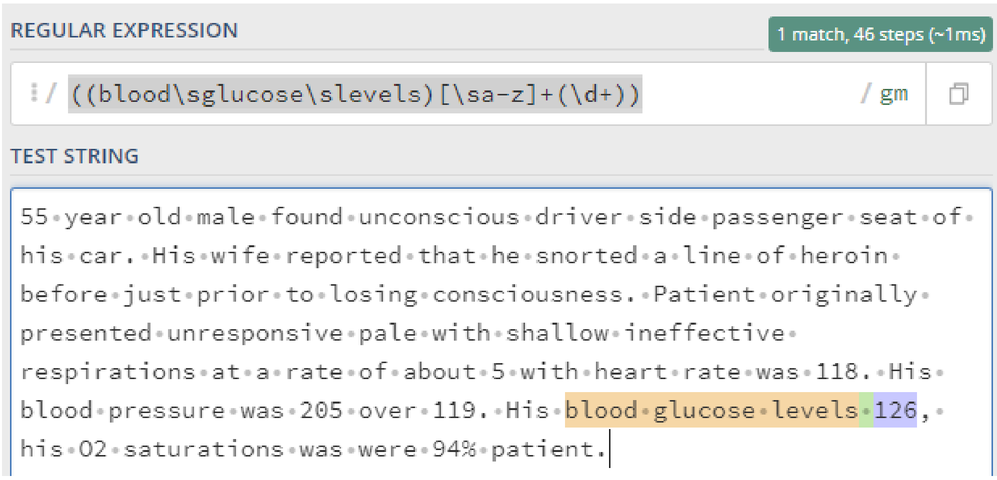


Figure 3.11. Sample Regex and Matched Output [42]

Figure 3.11 provides an example of a regex pattern and the matched output in test string is highlighted based on grouping.

3.1.5 Syntactic Pattern Matcher

While regular expressions help us to match patterns in text, spacy's pipeline component EntityRuler provides a rule based matcher that helps to identify tokens, entities and phrases based on pattern dictionaries[43]. Each pattern dictionary contains a "label" key which provides the label to be assigned to the entity if the pattern is matched, and a "pattern" key which specifies the match pattern in simple text form or as regular expressions[43]. The two types of accepted patters are:

- Phrase patterns to match exact string.
  Example, the gender of a male patient mentioned in a text can be
  retrieved as {"label": "GENDER", "pattern": "male" }

- Token patterns where a list of pattern values represent a single token
  Example, {"label": "BP",
  "pattern": [{"TEXT" : {"REGEX": "(\d+)"}},
  {"TEXT" : {"REGEX": "(over—\ /—-)"}}, {"TEXT" : {"REGEX": "(\d+)"}}]}

When the nlp object with EntityRuler as an added pipeline component is called on a text, pattern matches in the text are retrieved as entities with the corresponding label as the entity name. The first pattern matching the highest number of tokens takes priority when overlapping matches occur[43]. When the text contains more than one sentence, a sentencizer is used to separate sentences based on custom sentence boundary detection logic.

3.1.6 Medical Name Entity Recognition

Med7 is a refined named-entity recognition model trained mainly to identify seven categories: drug names, route, frequency, dosage, strength, form and duration[44]. SpaCy's implementation of a cloze-style word reconstruction, similar to the masked language model objectives introduced in BERT is used to predict word's vector using a static embedding table with a cosine loss function[37]. Information regarding treatment provided to patient as well as medical drugs taken by patient can be identified using Med7.

3.2 Medical Narrative Classification

Medical narratives obtained from the first responders at an incident site describes different mechanisms of injury. Using natural language processing techniques and machine learning, the mechanism of injury can be retrieved and reported. Figure 3.12 describes 4 classification models used for multi-class categorization to determine the mechanism of injury from a narrative text.

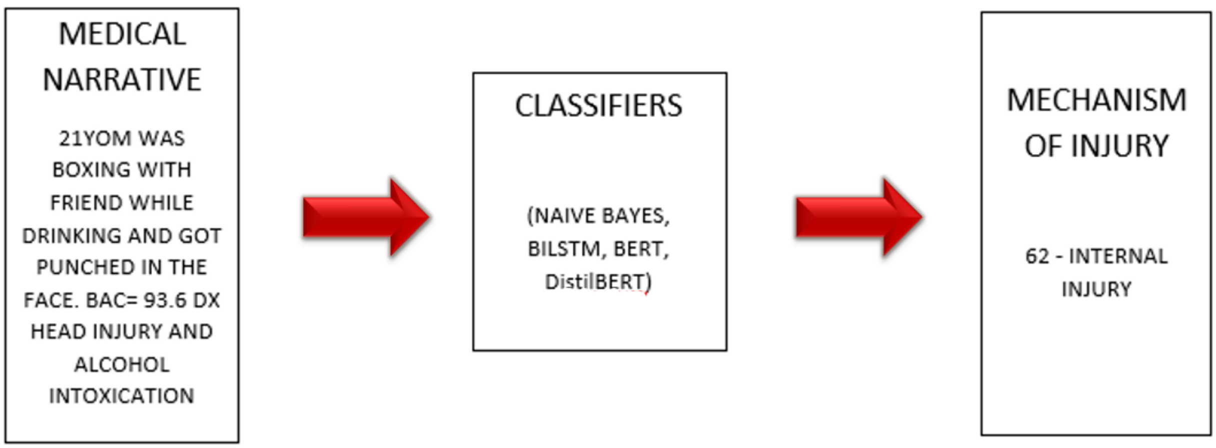


Figure 3.12. Medical Narrative Classification

Natural Language Processing is the field of artificial intelligence where computers are programmed to learn and analyze human languages using large volumes of data. NLP can be used to classify textual data into multiple classes based on the contextual information specified in the text. This paper compares some of the most common techniques available to classify EMS data based on the mechanism of injury described in the narrative.

3.2.1 Naive Bayes Classifier

The Bag-of-Words model is used to represent the narrative text as a bag of unique words disregarding the grammatical structure and order of occurrence but keeping the frequency of appearance. Each word in the vocabulary(bag of words) becomes a feature and a document is represented by a vector with the same length as the vocabulary. Terms frequency is often not the best representation for text as the corpus contains common words like 'the, is, a' with high frequency but little predictive power over the target variable. To address this problem, there is an advanced variant of the Bag-of-Words that uses the term frequency–inverse document frequency (or Tf–Idf) where the value of a word increases proportionally to count, but it is inversely proportional to the frequency of the word in the corpus[45]. Using the process of feature selection, the irrelevant features or words can be discarded hence reducing the dimensionality. Each output class is then described by a set of words that most likely represent itself. For example, top features from the set of medical narrative texts that represent '73 - RADIATION' mechanism of injury are 'burn, eye, pool, swimming, wearing, wo, 2nd, developed, yesterday, day'.

```
# 65 - ANOXIA:
  . selected features: 254
  . top features: caught,co,fd,fell,fire,found,headache,home,house,present

# 66 - HEMORRHAGE:
  . selected features: 99
  . top features: eye,face,hit face,nose,started,toilet,nasal,shower,trauma,taking

# 67 - ELECTRIC SHOCK:
  . selected features: 40
  . top features: felt,hand,put,stuck,water,working,body,got,hand dx,burn

# 68 - POISONING:
  . selected features: 435
  . top features: accidentally,ankle,back,bag,bottle,child,cleaning,contusion,drinking,dx

# 69 - SUBMERSION:
  . selected features: 105
  . top features: bathtub,came,found,jumped,mom,pool,pulled,started,swallowed,swimming

# 71 - OTHER:
  . selected features: 436
  . top features: 2nd,5th,abrasion,acc,accidentally,acute,admitted,ago,ago dx,ankle

# 72 - AVULSION:
  . selected features: 308
  . top features: 4th,avulsion,caught,cut,cut finger,cutting,door,door dx,elbow,finger
```

Figure 3.13. Top selected features for some mechanisms of injury

This model is then used to train a Naive Bayes machine learning model that works on the Bayes Theorem to classify the text into the set of desired multi-class output. Bayes' Theorem is a rule that uses probability to make predictions based on prior knowledge of conditions that might be related[45]. It is defined by the formula

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

where $P(A \mid B)$ is the probability of occurrence of event A given that event B has occurred, $P(B \mid A)$ is the probability of occurrence of event B given that event A has occurred and the probabilities of the individual events A and B are given by P(A) and P(B) respectively.

3.2.2 BiLSTM Classifier

Secondly, the word2vec model is used along with BiLSTM classifier. Word Embedding is the collective name for feature learning techniques where words from the vocabulary are mapped to vectors of real numbers calculated from the probability distribution for each word appearing before or after another so that words of the same context that usually appear together in the corpus will be close in the vector space as well[45].

sustained	rt	hand	crush	injury	hitting	hand	splitting	maul						
57	34	18	569	14	42	18	3524	15647	0	0	0	0	0	0

Figure 3.14. Sentence representation using word indices

Word2Vec models are generated through 2 main approaches, Skip-gram or Continuous Bag of Words. In Skip-gram model, a single word can be used to predict the context and in continuous bag of words (BOW), a set of words in a specific context is used to predict a single word. Using word embedding, the medical narrative texts are represented by a fixed dimension padded vector where each word is represented by its word index in the complete vocabulary set and given as input to a bidirectional long short term memory(BiLSTM) network.

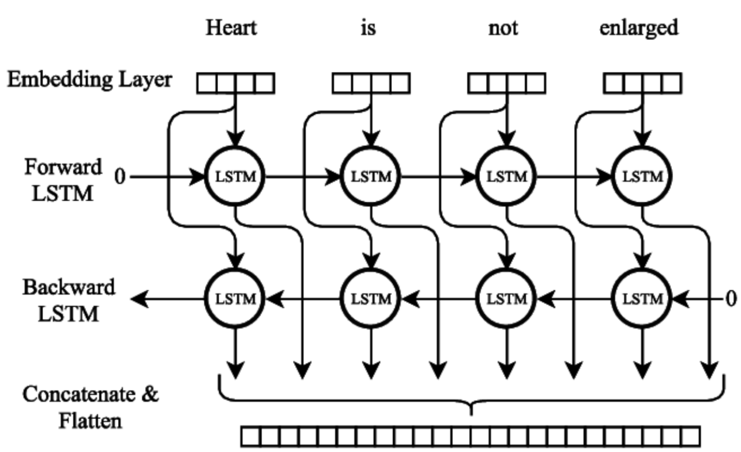


Figure 3.15. BiLSTM Model[46]

Padding and truncation ensures the length of vectors are uniform by inserting 0 as padding element if the size of the sentence is smaller than the defined vector dimension and truncating the sentences into multiple vectors if they are too long. Fig3.14 shows an example of a sentence of length 9 being represented by a padded vector of size 15.

The Bidirectional LSTM, or biLSTM is a type of recurrent neural network that consists of two LSTMs where the input vector is processed in both forward and backward direction effectively increasing the amount of information available to the network and consequently improving the context available to the algorithm[46]. It also makes use of the known property of RNN where previous processed sequential information is tracked in memory and used in future prediction or classification tasks.

3.2.3 BERT Classifiers

Since word embeddings are applied in a context-free manner, the same word that could have two different meanings based on the context in which it appears is represented by the same vector. To resolve this problem of polysemous disambiguation[45], language models or contextualized dynamic word embeddings have been developed.

Google's Bidirectional Encoder Representations from Transformers(BERT)[22] is a bidirectional model trained on a very large amount of unlabeled text that combines ELMO's context embedding which uses BiLSTM to assign embedding to a word based on the entire sentence, and several Transformers that use attention mechanisms to improve performance[45]. Since the vector BERT assigns to a word is a function of the entire sentence, a word can have different vectors based on the contexts[45].

The BERT can either be trained from scratch with a specific large corpus and then be used as a classifier or the word embeddings from the trained BERT model can be extracted and then used in an embedding layer for classification or the pre-trained model provided by the developers could be reused and fine-tuned to the specific requirement at hand also known as transfer learning[45]. Since training the mode from scratch requires an extensive amount of time and could cause problems of overfitting, this paper fine-tunes the pre-trained model of both the lighter version trained on about 66 million parameters called Distil-BERT and the actual version trained on a larger number of parameters ranging from 110 to 300 million for the purpose of classification of medical narratives.

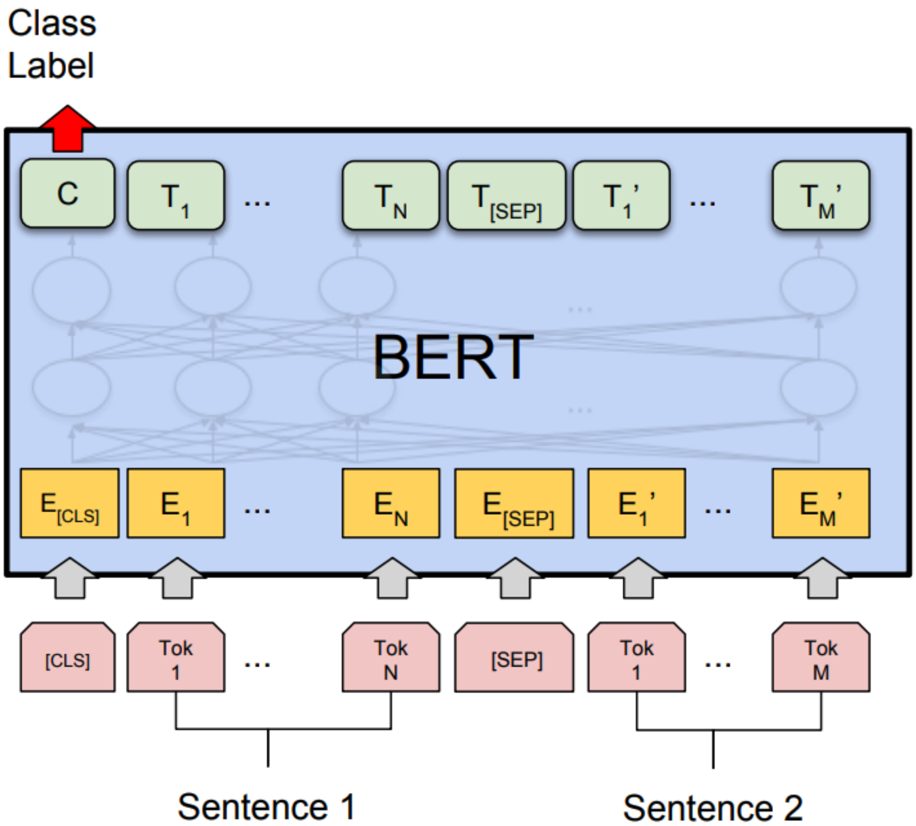


Figure 3.16. BERT Model[47]

As part of the process of fine-tuning, the input text is tokenized and represented by 3 vectors.

- Token Ids which are numerical representations of the word tokens building the sequences including tokens like [CS] to represent the start of a sentence, [SEP] to mark the separation between sentences and [PAD] to denote the padding element.
- Masks are used to distinguish between textual words and padding elements.
- Segment keeps track of the number of [SEP] tokens encountered.

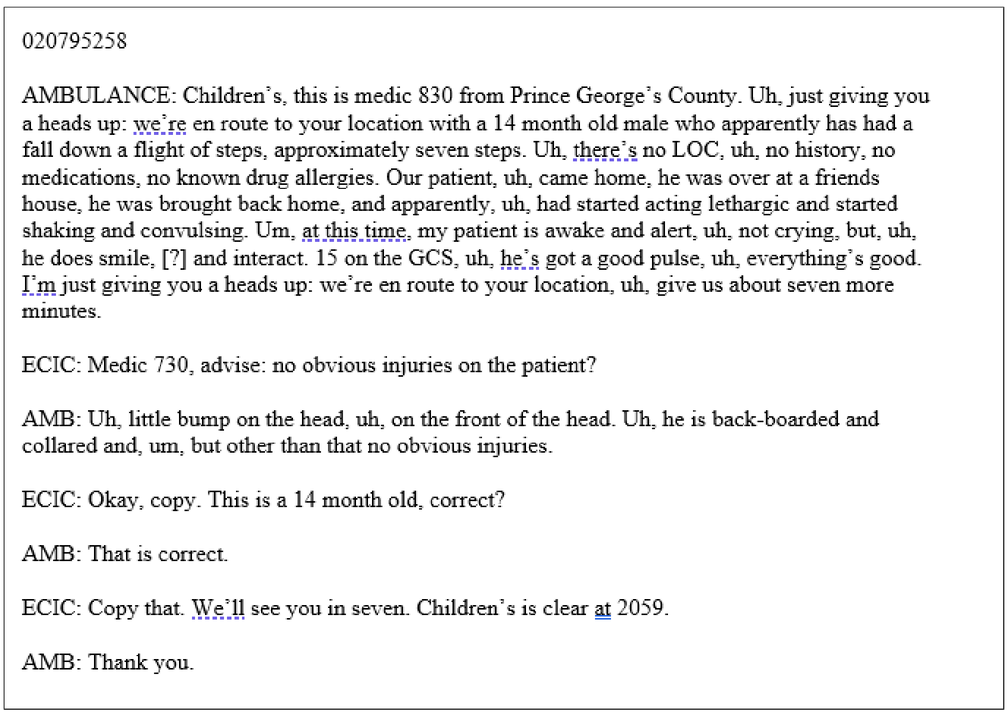
Since BERT splits unknown words into sub-tokens until it finds known unigrams, the maximum sequence length chosen is larger compared to the word2vec model. The deep learning model with transfer learning from the pre-trained BERT is built and run on the medical narratives to classify it and then performance was evaluated.

4. EXPERIMENTAL SETUP

4.1 Information Extraction

As can be seen from the information provided in the methodology section, medical word documents are processed to automatically fill an EMR form with the help of natural language processing and syntactic analysis. The details regarding the dataset used and code configurations to implement the solution are described in this section.

4.1.1 Dataset



020795258

AMBULANCE: Children's, this is medic 830 from Prince George's County. Uh, just giving you a heads up: we're en route to your location with a 14 month old male who apparently has had a fall down a flight of steps, approximately seven steps. Uh, there's no LOC, uh, no history, no medications, no known drug allergies. Our patient, uh, came home, he was over at a friends house, he was brought back home, and apparently, uh, had started acting lethargic and started shaking and convulsing. Um, at this time, my patient is awake and alert, uh, not crying, but, uh, he does smile, [?] and interact. 15 on the GCS, uh, he's got a good pulse, uh, everything's good. I'm just giving you a heads up: we're en route to your location, uh, give us about seven more minutes.

ECIC: Medic 730, advise: no obvious injuries on the patient?

AMB: Uh, little bump on the head, uh, on the front of the head. Uh, he is back-boarded and collared and, um, but other than that no obvious injuries.

ECIC: Okay, copy. This is a 14 month old, correct?

AMB: That is correct.

ECIC: Copy that. We'll see you in seven. Children's is clear at 2059.

AMB: Thank you.

Figure 4.1. Sample Transcript

The medical data used for the purpose of this research paper was a set of 60+ documents captured over a four-month period(June, 2009 - September, 2009) obtained from the recorded communications(verbal reports/notification calls) between EMS staff and a dispatcher or staff at the emergency department of Children's National Medical Center(CNMC) for trauma calls. The length of communications varied from 30 seconds to 4 minutes, and depended on the available information about the patient, the number of incoming patients and the purpose of the call. For example, a call that describe details of a critical patient as the ambulance is travelling to the hospital is much longer than a call where the EMS staff is providing only information regarding number of patients and estimated time of arrival based on location. The documents were then processed to retain information specific to patient similar to the scenario when the voice of an EMS staff, reporting patient information at an emergency site, is being recorded by an voice recognition system. With an exception of a few documents, the information in these medical narratives contained details regarding the complaint raised, vitals of patient, cause of injury, patient history and treatment provided by EMS staff.

4.1.2 Code Configurations

The medical narratives are then further processed through several stages of extraction in code to obtain a set of 24 field names and it's corresponding values from the text.

Initially, the publicly available biomedical concept extractor MetaMap was installed following setup instructions provided by Willie Rogers at the United States Library of Medicine(nlm) website[17]. To control Metamap's internal behavior, the following processing options were set.

- -y(- -word_sense_disambiguation) : causes metamap to disambiguate among concepts that score equally well in matching input text[17],
- -u(- -unique_acros_abbrs_only) : causes metamap to restrict the generation of acronym/abbreviation (AA) variants to those forms with unique expansions[17].
- - -negex : causes metamap to output a list of negated umls concepts occurring in the input and the associated strings that caused the negation[17]. This option is very useful in medical domain as medical professionals are trained to provide the salient features of a patient in a short and precise manner negating possible ambiguity in language.
- - -conj: Causes MetaMap's phrase chunker to recombine smaller phrases separated by a conjunction[17].

Of the 127 semantic types available in MetaMap, ten specific types were chosen as they specifically reflect the various findings of In-field EMS personnel in their preliminary diagnosis. They are as follows: *'Body Location or Region', 'Body Part, Organ, or Organ Component', 'Clinical Attribute','Clinical Drug', 'Disease or Syndrome', 'Finding', 'Injury or Poisoning', 'Medical Device', 'Pathologic Function'* and *'Sign or Symptom'*. Syntactic dependency parser like spacy was then used to find the syntactic relations between words from the filtered set as well as from the raw text to ensure relevant information was not missed. Significant data that was still not captured was extracted using regular expressions to match reoccurring patterns and to format final data. Concepts that contain number values are extracted from the narrative text using simple regular expressions. The final extracted data was populated for 24 concept fields similar to that seen on an EMR form. Table 4.1 describes what each of the 24 concept fields represent.

4.1.3 Concept Extraction

AGE: While age of a patient is best calculated based on date of birth, since this research uses data where personally identifiable information has been removed the value is retrieved based on the calculation of the EMS staff mentioned in the narrative text in terms or years or months.

GENDER: Gender of the patient is either male or female extracted using spacy's entity ruler constraint.

BP: After retrieving data regarding blood pressure using regular expression, a check is performed to ensure that the systolic blood pressure(first number value) mentioned is always greater than the diastolic blood pressure(second number value) with an exception of 0/0 if the patient is dead.

PULSE: Information regarding pulse stated by EMS staff usually contain 2 parts where the first part is a number value and the second part states information about quality of pulse like regular, irregular or good. As pulse value is often found in the narrative as either heart rate or pulse, simple regex is used to extract the concept.

Table 4.1. Concept fields and its description

Concept Field	Description
AGE	Age
GENDER	Gender
BP	Blood pressure
PULSE	Heart rate or pulse
GCS	Glasgow Comma Scale(3 - 15)
TRAUMA SCALE	Trauma Priority value
RESP	Respiration rate
B.G.L	Blood glucose level
SPO2	Oxygen saturation
MENTAL ST	Mental status
PATIENT COND	Patient condition
CURRENT MEDICATION	Current medications taken by patient
ALLERGIES	Patient allergies
PAST MEDICAL HISTORY	Past medical history of patient
PAIN	Descriptions about pain
PUPILS	Description about both left and right pupils
LUNG SOUNDS	Description for both left and right lung
VERBAL	Description about speech of patient
AIRWAY	Details regarding airway tract
INJURY	Trauma, wounds or any other injury
MECHANISM OF INJURY	Mechanism of injury
COMPLAINT	Complaint from patient
TREATMENT	Treatment provided by EMS staff
NOTES	Complete description of patient case

TRAUMA SCALE: Since the dataset used for this research contained mostly trauma calls, the trauma priority level defined by the dispatcher or EMT is identified as trauma scale with the help of syntactic dependency parser spacy that searches for dependent words related to priority unlike the trauma scale in usual EMR forms which refers to information that contain a number score ranging from 0 to 12 and a description of the severity of trauma.

REP: Besides containing the respiration rate value which will always be a positive number in an EMR form, respirations could also refer to the quality of breathing of the patient indicated by phrases like labored or shallow breathing. The number value for respirations was extracted using regular expressions. For example, the medical transcript *"A 14 year old male, one GSW to the right hip. no exit. We have got a blood pressure of 116 over 70, heart rate 90, respiration's about a 20 right now, ALS is, being provided at this time. He does have a line. He is fully boarded and collared. He is alert and oriented. Equal lung sounds bilaterally. You can actually see the bullet, right in the front pelvis. you can see it like in his front hip."* had extracted value for 'RESP' as 20.

B.G.L: Blood glucose levels(BGL) measured by a glucometer, helps to provide early and condition - specific treatment[48]. One of the many varieties of medical conditions that warrant prehospital blood glucose analysis is an altered mental status[48].

SPO2: With the help of a noninvasive pulse oximeter that measures both pulse rate and the arterial oxygen saturation of hemoglobin at the peripheral capillary level, an EMT can determine if a patient is hypoxic[49]. The normal range for SPO2 is between 94% to 99%.

GCS, VERBAL: The Glasgow Coma Scale was designed to assess the duration and depth of coma and impaired consciousness which helps to gauge the impact of a wide variety of conditions such as acute brain damage due to traumatic and/or vascular injuries or infections, metabolic disorders (e.g., hepatic or renal failure, hypoglycemia, diabetic ketosis), etc[50]. Figure 4.2 describes how the GSC score is calculated as defined by CDC.

Eye Opening Response

- Spontaneous--open with blinking at baseline **4 points**
- To verbal stimuli, command, speech **3 points**
- To pain only (not applied to face) **2 points**
- No response **1 point**

Verbal Response

- Oriented **5 points**
- Confused conversation, but able to answer questions **4 points**
- Inappropriate words **3 points**
- Incomprehensible speech **2 points**
- No response **1 point**

Motor Response

- Obeys commands for movement 6 points
- Purposeful movement to painful stimulus 5 points
- Withdraws in response to pain 4 points
- Flexion in response to pain (decorticate posturing) 3 points
- Extension response in response to pain (decerebrate posturing) 2 points
- No response 1 point

Figure 4.2. Glasgow Coma Scale[50]

Although it is expected that in the care of an individual patient, the ratings of the three criteria in the scale should be assessed, monitored, reported, and communicated separately, the transcripts analyzed mostly had only the final GCS score which was extracted from narrative using regex as a number value, and the verbal assessment description in the narrative was extracted as the 'verbal' concept with the help of spacy's syntactic dependency parser. For example, the GCS value '15' was extracted using pattern matching with regular expression from the medical transcript *"A 14 month old male who apparently has had a fall down a flight of steps, approximately seven steps. There's no LOC, no history, no medications, no known drug allergies. Our patient, came home, he was over at a friend's house, he was brought back home, and apparently, had started acting lethargic and started shaking and convulsing. At this time, my patient is awake and alert, not crying, but, he does smile, and interact. 15 on the GCS, he has got a good pulse, everything's good. Little bump on the head, on the front of the head. He is backboarded and collared and, but other than that no obvious injuries."*.

MENTAL ST: The mental status examination is a structured assessment of the patient's behavioral and cognitive functioning which includes descriptions of the patient's appearance and general behavior, level of consciousness and attentiveness, motor and speech activity, mood and affect, thought and perception, attitude and insight, and higher cognitive abilities of which the specific cognitive functions of alertness, language, memory, constructional ability, and abstract reasoning are of utmost priority in the rapid assessment done by the first responder[51]. From the narrative text, the 'mental status' concept was extracted from phrases retrieved by metamap's semantic type 'Finding' and then refining the data using spacy's dependency parser to retrieve related words that are modifiers or connected by conjunctions. For example, some of the phrases encountered were *alert and oriented(AnO), lethargic, confused, loosing interval, etc.* to describe the mental status of the patient.

PATIENT COND: Based on the initial inspection which considers patient vitals and other factors combined, the patient condition is categorized using the acronym CUPS where 'C' for critical used when the patient is either receiving CPR, in respiratory arrest, or requiring and receiving life sustaining ventilatory/circulatory support, 'U' for unstable used when the patient has a poor general impression being unresponsive with no gag or cough reflexes,

’P’ for potentially unstable when patient is responsive but unable to follow commands, has difficulty in breathing, has pale skin or other signs of poor perfusion (shock), is having a complicated childbirth situation, has uncontrolled bleeding, has severe pain in any part of the body especially severe chest pain with a systolic BP of less than 100mmHg or is unable to move any part of the body and, lastly, ’S’ for stable when patient has a minor illness, minor isolated injury, uncomplicated extremity injuries, and/or any other condition that cannot be categorized as Critical, Unstable or Potentially unstable[52]. This information is retrieved with the ’Finding’ semantic type of metamap sometimes as well as spacy’s syntactic analysis that search for the subject word and its related connection. For example, the value of ’PATIENT COND’ concept is ’patient is stable right now’ for the medical transcript *Notification of a Priority 2 transport. We have a 14 year old male who was hit by a bus, with a possible femur and pelvis fracture. He is complaining of pain in those regions. The, patient is stable right now. He has a blood pressure of 120 over 78, a pulse of 111, and a respiration, 16. Patient is in full, spinal immobilization and an IV has also been established. no loss of consciousness* using word search to extract the word stable and then analyzing with spacy’s dependency parser adding the constraint that it describes a pronoun like he/she or a noun like patient. With ’stable’ as the root word(checked by retrieving the parent/head of the token, and if the parent is same as the token, it is the root word), the token tag is checked to identify if it’s a noun(NN), verb(VB) or adjective(JJ) after which the token children are retrieved based on it’s dependency connection to the word ’stable’ being either a conjunction or modifier in this case. The retrieved text is then compared to the medical narrative text to obtain the missing adverbs, punctuation’s, etc and reorder the words to match the phrase in text.

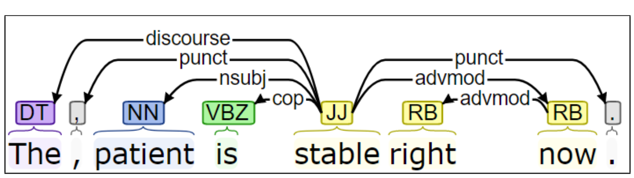


Figure 4.3. Dependency Tree Visualization for ’The, patient is stable right now’[38]

PAST MEDICAL HISTORY, CURRENT MEDICATIONS: The patient's past medical history refers to chronic health conditions or surgeries that the patient has had previously along with the current medications taken by the patient. This can provide important clinical insight when triaging, diagnosing, and treating patients with critical, emergent illness. In life-threatening emergencies when the patient is often not able to provide information, the medical professionals use other sources of information like family, other bystanders, and/or medications containers to construct a quick draft chronic health profile for the patient[53]. From the medical narrative, past medical history and medications concepts are retrieved using metamap semantic types 'Finding', 'Clinical Attribute', 'Clinical Drug' and 'Disease or Syndrome' combined with medicine extractor algorithm Med7 and spacy's dependency parser. The current medications taken by patient is distinguished from the medications provided by the EMS staff by using syntactic analysis to find information that specifies that the patient is taking the medicine compared to the patient being given the medication.

For example, the medical transcript *A 14 month old male who apparently has had a fall down a flight of steps, approximately seven steps. There's no LOC, no history, no medications, no known drug allergies. Our patient, came home, he was over at a friend's house, he was brought back home, and apparently, had started acting lethargic and started shaking and convulsing. At this time, my patient is awake and alert, not crying, but, he does smile, and interact. 15 on the GCS, he has got a good pulse, everything's good. Little bump on the head, on the front of the head. He is backboarded and collared and, but other than that no obvious injuries* has 'PATIENT MEDICAL HISTORY' as 'no history', ALLERGIES as 'no known drug allergies' with metamap output as shown below.

- Phrase: no history',

  Meta Mapping (1000):

  1000 N History (Medical History) [Finding]

- Phrase: '

- Phrase: no medications',

  Meta Mapping (1000):

  1000 Medications (Medications:-:Point in time:P̂atient:-) [Clinical Attribute]

- Phrase: ’

- Phrase: no known drug allergies’,

  Meta Mapping (901):

  660 Known [Qualitative Concept]

  901 N Drug Allergies (Drug Allergy) [Pathologic Function]

ALLERGIES: Allergies refer to medication that cause an anaphylactic reaction in the patient causing an upper airway obstruction often described by the lung sound description ’stridor’ heard during the inspiratory phase of the respiratory cycle. With metamap semantic types ’Finding’ and ’Pathologic Function’, phrases describing patient allergies in the narrative was retrieved and then refined using spacy.

LUNG SOUNDS: Lung sounds are of 2 types: expiration and inspiration sounds. The duration and clarity of lung sounds helps to identify many acute and chronic problems which may or may not be directly related to the lungs[54]. Lung sounds are described as either ’equal’, ’clear’, diminished’, ’wheezes’, ’rhonchi’, ’crackles(rales)’ or ’absent’[55]. The information regarding lung sounds in the narrative text is retrieved by first obtaining phrases related to lungs using metamap semantic type ’Body Part, Organ, or Organ Component’ and then words describing lungs are obtained using syntactic dependency parser that identifies adjectives and modifiers. The produced output is then filtered and reorder the combined words to reflect the same order of words in the text. For example, the medical transcript ”*A 14 year old male, one GSW to the right hip. No exit. We have got a blood pressure of 116 over 70, heart rate 90, respiration’s about a 20 right now, ALS is, being provided at this time. He does have a line. He is fully boarded and collared. He is alert and oriented. Equal lung sounds bilaterally. You can actually see the bullet, right in the front pelvis. You can see it like in his front hip*”, metamap retrieves the phrase ’equal lung sounds bilaterally’ as 3 separate phrases as shown below which is then processed by spacy.

- Phrase: ’Equal lung

  Meta Mapping (888):

  694 Equal (Relational Operator - Equal) [Intellectual Product]

  861 LUNG (Lung) [Body Part, Organ, or Organ Component]

- Phrase: sounds

  Meta Mapping (1000):

  1000 Sounds (Sound - physical agent) [Natural Phenomenon or Process]

- Phrase: bilaterally'

PUPILS: While most often analysis for pupils is distinguished as analysis for the left or right pupil, since the dataset mostly had scenarios where patient pupils description was combined for both eyes and stated as either equal or unequal pupils that is/are reactive or unresponsive to light, extracted concept field 'PUPILS' refer to both pupils. Figure

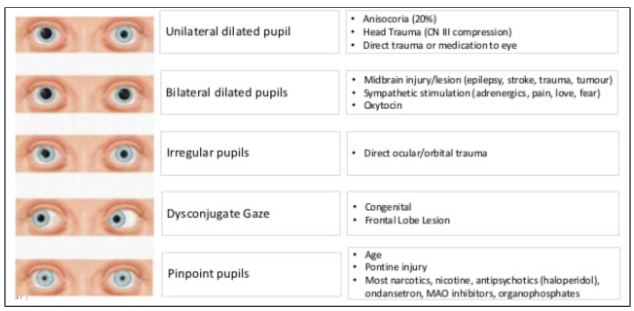


Figure 4.4. Eye Assessment[56]

4.3 provides some of the information regarding eye assessment performed by a paramedic. Concepts related to pupils is extracted first as phrases related to eyes or pupils by metamap semantic type 'Body Part, Organ, or Organ Component' and then using syntactic analysis, adjectives describing pupils and modifiers linked by conjunctions are retrieved from the complete sentence to get the expected mapping. For example, for the medical transcript '*A six year old female, victim of a fall. Right now she has pain in her head and on the bridge of her nose. Vitals are stable. 100% blood pressure palp, pupils are equal and reactive to light. 104 pulse. 99% O-2 stat. We have her on blow-by oxygen right now. Transporting Priority 2. Mother is claiming that she had some sort of convulsion after the table fell on her head. Her mental status now completely alert and oriented.*' the value of 'PUPILS' after processing using a similar algorithm as that used for retrieving patient condition is 'pupils are equal and reactive to light'. The metamap output is shown below.

- Phrase: ’pupils

  Meta Mapping (1000):

  1000 Pupils (Pupil) [Body Part, Organ, or Organ Component]

- Phrase: equal

  Meta Mapping (1000):

  1000 Equal [Qualitative Concept]

- Phrase: reactive to light’,

  Meta Mapping (1000):

  1000 Reactive to light [Finding]

INJURY, MECHANISM OF INJURY: Injury information was extracted using metamap semantic type ’Injury or Poisoning’ combined with phrase extraction using commonly used words related to injury like scratch, bump, bruise, etc. to capture information missed by metamap. Mechanism of injury usually presents the problem for trauma calls while nature of illness is used in the case of medical calls related to diseases or illness. Since metamap algorithm did not distinguish between injury and mechanism of injury, with the help of syntactic analysis and simple word search, sentences that contained description of the mechanism of injury was extracted. Words commonly related to mechanism of injury that are used for this study include ’fall’, ’gunshot’, ’struck’, ’fire’, ’attack’, ’collision’, ’assault’, ’stab’, ’hit’, ’suicide’, ’drowning’, ’crash’ and ’GSW’(abbreviation for gunshot wound).

COMPLAINT: Complaint usually contains information described by the patient but in cases where patient says that they do not have a complaint and still needs the help of the emergency medical service, the description provided by the first responder is considered as complaint.

TREATMENT: The medication and treatment provided by the EMS staff at the site or during transportation to the hospital is considered in general as ’TREATMENT’. The concept is extracted using information retrieved by metamap semantic types ’Clinical Attribute’,’Clinical Drug’ and ’Medical Device’ along with phrase extraction based on certain words and spacy’s syntactic analysis.

For example, the medical transcript "*A 14 month old male who apparently has had a fall down a flight of steps, approximately seven steps. There's no LOC, no history, no medications, no known drug allergies. Our patient, came home, he was over at a friend's house, he was brought back home, and apparently, had started acting lethargic and started shaking and convulsing. At this time, my patient is awake and alert, not crying, but, he does smile, and interact. 15 on the GCS, he has got a good pulse, everything's good. Little bump on the head, on the front of the head. He is backboarded and collared and, but other than that no obvious injuries.*" is processed to obtain an output with 24 concept fields and it corresponding values from the narrative as shown in Figure 4.3.

```
{
    "GENDER": [
        "male"
    ],
    "AGE": [
        "14 month old"
    ],
    "GCS": [
        "15"
    ],
    "PULSE": [
        ""
    ],
    "BP": [
        ""
    ],
    "RESP": [
        ""
    ],
    "B.G.L": [
        ""
    ],
    "SPO2": [
        ""
    ],
    "MENTAL ST": [
        "not crying",
        "awake",
        "alert",
        "no LOC",
        "had started acting lethargic and started shaking and
convulsing"
    ],
    "PATIENT COND": [
        ""
    ],
    "CURRENT MEDICATION": [
        ""
    ],
    "ALLERGIES": [
        "no known drug allergies"

    ],
    "PAST MEDICAL HISTORY": [
        "no history"
    ],
    "PAIN": [
        ""
    ],
    "TRAUMA SCALE": [
        ""
    ],
    "PUPILS": [
        ""
    ],
    "LUNG SOUNDS": [
        ""
    ],
    "VERBAL": [
        ""
    ],
    "AIRWAY": [
        ""
    ],
    "INJURY": [
        "but other than that no obvious injuries",
        "Little bump on the head"
    ],
    "MECHANISM OF INJURY": [
        "A 14 month old male who apparently has had a fall down a
flight of steps"
    ],
    "COMPLAINT": [
        ""
    ],
    "TREATMENT": [
        "backboarded",
        "collared"
    ],
    "NOTES": [
        "A 14 month old male who apparently has had a fall down a
flight of steps, approximately seven steps. there’s no LOC, no
history, no medications, no known drug allergies. Our patient,
came home, he was over at a friends house, he was brought back
home, and apparently, had started acting lethargic and started
shaking and convulsing. at this time, my patient is awake and
alert, not crying, but, he does smile, and interact. 15 on the
GCS, he’s got a good pulse, everything’s good. Little bump on the
head, on the front of the head. he is backboarded and collared
and, but other than that no obvious injuries."
    ]
}
```


Figure 4.5. JSON Output

4.2 Classification

While extracting concepts from medical narratives, information related to one of the fields, namely 'MECHANISM OF INJURY' was often too long and had some extra information. To resolve that issue, 3 types of classification models using a much larger dataset was trained and tested for classification accuracy, the implementation details of which is described further in this section.

4.2.1 Dataset

Data used for classification was obtained by querying the National Electronic Injury Surveillance System (NEISS) for a range of 3 years starting from 2017 until 2019. NEISS is an injury surveillance and follow-back system that provides timely data on consumer product-related injuries occurring in the United States.

Table 4.2. Sample NEISS Data used for Classification

TEXT	Y
21YOM WAS BOXING WITH FRIEND WHILE DRINKING AND GOT PUNCHED IN THE FACE. BAC= 93.6 DX HEAD INJURY AND ALCOHOL INTOXICATION	62-INTERNAL INJURY
71YOM FELL ON FLOOR IN BATHROOM. DX RIB CONTUSION	53-CONTUSIONS, ABR.
16YOM TRIPPED OVER A RUG AND FELL INTO A BURNING FIREPLACEDX BURNS INVOLVING LESS THAN 10% OF BODY SURFACE, FOREARM	51-BURNS, THERMAL
58YOF FX LWR LEG- TRIP ON DOG, FELL PORCH STEPS- ETOH- XFER DUKE	57-FRACTURE
22YOM D'LOC KNEE REPOSITIONING SELF IN BED	55-DISLOCATION
31YOF CONCUSSION- WALKED UP STEPS, FELL FLOOR- ETOH	52-CONCUSSION.
62YOF LAC LWR ARM- FELL AGAINST CHAIR	59-LACERATION

Injury data is collected from a sample of 100 U.S. hospital emergency departments selected from all 5,300+ hospitals in the United States with an emergency department[57]. The data contains 25 columns with values CPSC_Case_Number, Treatment_Date, Age, Sex, Race, Other_Race, Hispanic Origin, Body_Part, Diagnosis, Other_Diagnosis, Body_Part_2, Diagnosis_2, Other_Diagnosis_2, Disposition, Location, Fire_Involvement, Product_1, Product_2, Product_3, Alcohol involved, Drug involved, Narrative, Stratum, PSU and Weight with more than 300,000 records. For this study, data from the 'NARRATIVE' column was used as input to provide an output that specifies the mechanism of injury or diagnosis category. The data contains 30 different categories for mechanism of injury.

4.2.2 Code Configurations

Using the python nltk library, the data in the narrative column is pre-processed and cleaned by removing stop words and setting some of the words to its original form, also called lemmatization. The dataset is then partitioned as train and test data with 70% as train data and remaining 30% as test data. Firstly, for classification using Naive Bayes algorithm, the input data is represented by a feature matrix with the help of python library sklearn's TfidfVectorizer[58]. The number of rows in the feature matrix is equal to the number of records in training data and the number of columns is equal to the vocabulary size of the data, which is the count of unique words in the complete document. Treating each output category as binary values, a Chi-square test is performed to find the most relevant features from the narrative texts for each category and hence reducing the size of columns in the feature matrix which is then used as input for running the multinomial naive bayes classifier from the scikit-learn library[45]. Table 4.3 shows the predicted results for 10 samples and its corresponding probability.

Table 4.3. Naive Bayes Classification Results

True Value	Predicted Value	Probability
71-OTHER	71-OTHER	0.23
52-CONCUSSION	62-INTERNAL INJURY	0.44
53-CONTUSIONS,ABR.	59-LACERATION	0.4

Secondly, with the help of gensim python library, a skip-gram word2vec model is used to generate word vectors of size or dimension as 300 for each individual word in the input text. The narrative text is then transformed to a padded sequence of input index ids to get a feature matrix. The corresponding embedding matrix contains the word vectors linked to

Table 4.4. BiLSTM Classification Results

True Value	Predicted Value	Probability
71 - OTHER	71 - OTHER	0.85
52 - CONCUSSION	52 - CONCUSSION	1.0
53 - CONTUSIONS, ABR.	53 - CONTUSIONS, ABR.	1.0

each input index. Using the feature matrix as an embedding layer and the word vectors as weights, a BiLSTM neural network combined with attention mechanisms and dense layers to improve accuracy and predict probability of each output category for the output, is trained for 10 epochs and then tested on test data a few of which is shown in Table 4.4[45].

Table 4.5. Language model Classification Results

True Value	Distil-BERT Predicted Value	Prob	BERT Predicted Value	Prob
71-OTHER	71-OTHER	0.29	71-OTHER	0.27
52-CONCUSSION	59-LACERATION	0.4	62-INTERNAL INJURY	0.38
53-CONTUSIONS,ABR.	59-LACERATION	0.26	57-FRACTURE	0.33

Lastly, due to limitations on time and data size, pre-trained embeddings of google's language model BERT and its lighter version DistilBERT is fine-tuned instead of training a model from scratch. Feature engineering is performed on the input narrative text to which some special tokens are added to denote start and end of sentence then an id vector, mask vector and segment vector is generated for each input narrative text. The resulting feature matrix obtained has number of rows equal to the number of input narrative records with each row having 3 vectors. As part of building the deep learning model with transfer learning, the feature matrix is provided as input along with 2 additional dense layers added to predict probability of output after which the model is trained and tested.

5. EVALUATION

5.1 Information Extraction

The performance of the engineered algorithm was analyzed with the help of 2 evaluators, with one of them being the director of the Emergency Medical Services department having 30+ years of experience in EMS and the other evaluator being the Chief Data Officer of The Paramedic Foundation, president of American Paramedic Association and treasurer of the National EMS Management Association working on his PhD in Public Policy & Admission with an MS, NRP and FACPE degree.

FIELD	VALUE	MAPPING CORRECT?YES/NO	CORRECT MAPPING VALUE	TEXT
GENDER	['male']	YES		['A 14 month old male who apparently has had a fall down a flight of steps, approximately seven steps. there's no LOC, no history, no medications, no known drug allergies. Our patient, came home, he was over at a friends house, he was brought back home, and apparently, had started acting lethargic and started shaking and convulsing. at this time, my patient is awake and alert, not crying, but, he does smile, and interact. 15 on the GCS, he's got a good pulse, everything's good. Little bump on the head, on the front of the head. he is backboarded and collared and, but other than that no obvious injuries.']
AGE	['14 month old']	YES		
GCS	['15']	YES		
PULSE	[]	NO	"Got a good pulse"	
BP	[]	YES		
RESP	[]	YES		
B.G.L	[]	YES		
SPO2	[]	YES		
MENTAL ST	['not crying', 'awake', 'alert', 'no LOC', 'had started acting lethargic and started shaking and convulsing']	NO	patient is awake and alert, not crying, but he does smile, and interact	

Figure 5.1. Evaluation of Sample Document

Based on the evaluation of a set of 19 transcript outputs containing extracted/mapped concepts, a strict score and partial score was calculated for each extracted concept value from the narrative.

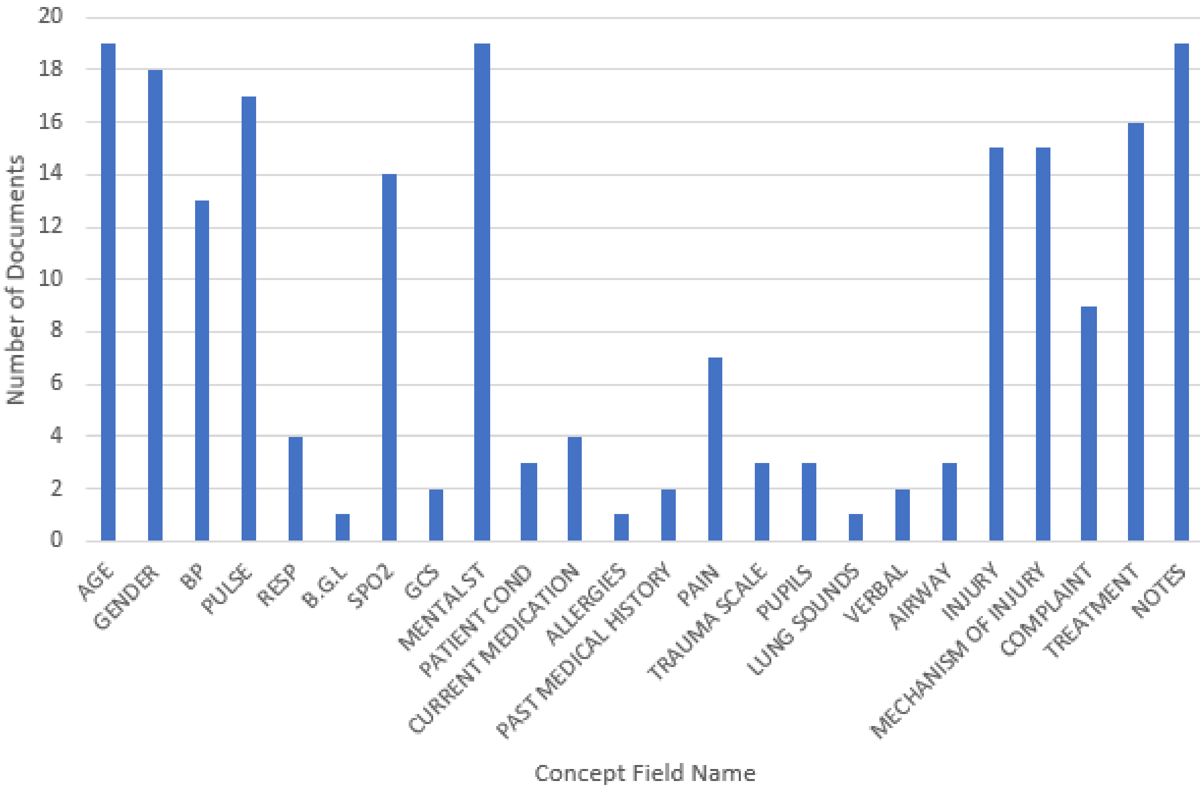


Figure 5.2. Concept Distribution for 19 documents

In Table 5.1, for each concept field, the strict score represents the ratio of number of correctly mapped documents to the total number of mapped documents and partial scores was calculated as the ratio of the number of documents with either partial or correctly mapped documents to the total number of documents. For example, in the medical transcript *A 14 year old male, who got out of bed and had some sort of syncopal episode, and took a fall striking his head. Says he then got up and tried to walk again and took two more falls. When they found him he was very disoriented, extremely diaphoretic, cold and clammy. Upon the arrival of EMS, said they thought he had a GCS of 14 and was confused. He had unequal pupils on the scene, 7 millimeters and 4 to 5 millimeters and they were extremely sluggish. Upon my arrival, the patient had a GCS of 15, airway is patent, breathing bit labored, no obvious external signs of trauma. Pupils are equal, round, and reactive at about 2 to 3 millimeters at this time. Current monitoring of vital signs, blood pressure is 118 over 69, sats 100 percent with a nasal cannula. He is secured on a backboard. When I arrived on the*

Table 5.1. Evaluation Score

FIELD	STRICT SCORE	PARTIAL SCORE
GENDER	1	1
AGE	1	1
BP	0.895	0.947
PULSE	0.947	0.947
SPO2	1	1
RESP	0.737	0.789
B.G.L	1	1
GCS	0.947	1
MENTAL ST	0.684	0.947
PATIENT COND	0.263	0.263
CURRENT MEDICATION	0.789	0.789
ALLERGIES	1	1
PAST MEDICAL HISTORY	0.947	0.947
PAIN	0.842	0.842
TRAUMA SCALE	0.842	1
PUPILS	0.947	0.947
LUNG SOUNDS	0.842	0.842
VERBAL	0.789	0.842
AIRWAY	0.947	1
INJURY	0.737	0.789
MECHANISM OF INJURY	0.474	0.737
COMPLAINT	0.842	0.895
TREATMENT	0.737	0.842
NOTES	1	1

scene he was still breathing hard as day, he was running, his heart rate was about 50. He has popped up from that, not sure if we have a medical? But, no real obvious signs of trauma, *and, no blood or fluids in his ears or his nose*, 'AGE' value '14 year old' is correct mapping whereas the mapped value for concept 'VERBAL' is 'confused' which is partially correct as the expected mapping was 'disoriented, confused'. While calculating the initial score, a mapping is considered partially correct if at least one word in current mapping matches the expected mapping. Similarly each concept field value is analyzed to be either correct or partially correct for each medical transcript after which the sum of total number of correct mappings and total number of partial mappings is calculated. The total number of current mappings for each concept is the number of total number of documents with mapping value for each concept that could be either empty if there is no mention of the concept in the narrative text.

$$\text{Average Strict Score} = \frac{\text{Number of Correct Mappings}}{\text{Total Number of documents}}$$

$$\text{Average Partial Score} = \frac{\text{Number of Partly Correct Mappings}}{\text{Total Number of documents}}$$

For 2 given input strings, Levenshtein Distance named after the Russian Scientist Vladimir Levenshtein, measures the minimum number of insertions, deletions or substitutions needed to change a one string into the other. In the case of partially correct mappings, the Fuzzy-Wuzzy library that works on the principle of Levenshtein Distance was used for measuring the similarity between the extracted value and ground truth.For analyzing the partially correct concept mappings, we calculate the fuzzywuzzy partial_ratio,token_sort_ratio and token_sort_ratio. The fuzzywuzzy ratio is a measure of the string similarity with a value in range [0,100]. With T as the total number of characters in both strings and M as the number of matches in the two string, the formula for calculation the ratio is

$$\text{Fuzzywuzzy Ratio} = \text{i}nt(round(\frac{2 * M * 100}{T}))$$

The normalized value of the fuzzywuzzy ratio is considered as the fuzzywuzzy score. The maximum value of the fuzzy wuzzy ratio similarity measures between the shorter string and every substring of the longer string gives the fuzzywuzzy partial ratio value. It is similar to the *contains* method available for string in most programming languages but it fails if

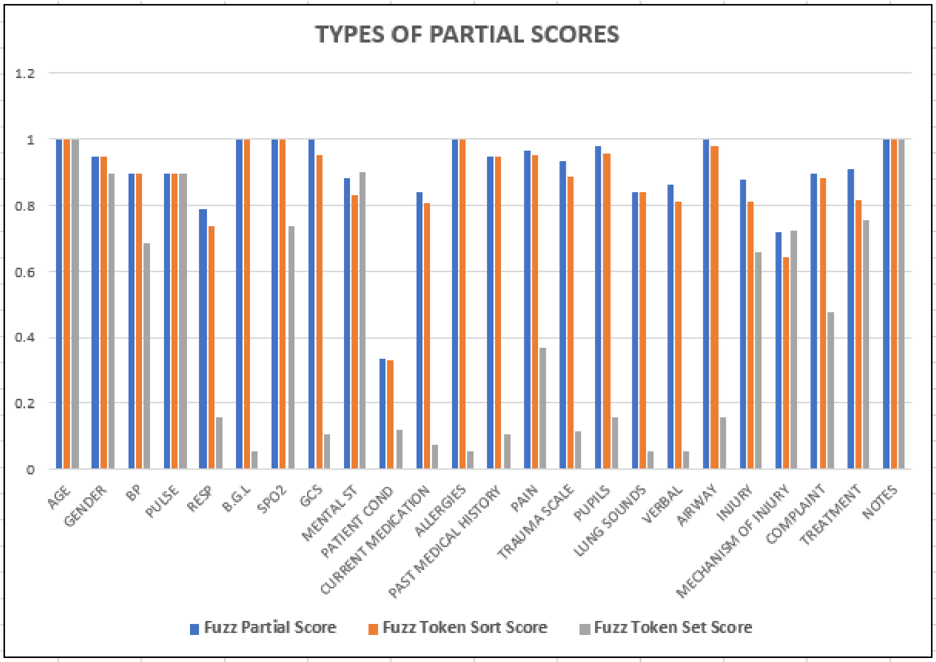


Figure 5.3. Average FuzzyWuzzy Scores

the order of words is not the same in the provided strings. This problem is overcome by the token_sort_ratio that splits the words or characters in both strings based on a delimiter, reorders both string and then compares the 2 strings. The token_set_ratio works similar to the token_sort_ratio with the additional benefit of removing duplicate substrings within a string before comparison. Figure 5.2 shows the average normalized fuzzywuzzy ratios for each concept field based on the evaluator's review.

$$\text{Average Fuzzywuzzy Score} = \frac{\text{Sum of Nomalized Fuzzywuzzy Ratios}}{\text{Total number of documents}}$$

5.2 Classification

While it is common to consider the best model as the one that predicts the maximum number of correct samples, the quality of outcome and time required to obtain the results matter when it comes to classification models. When analyzing the quality of the classification predictions, there are four categories to be considered.

- True Positives(TP): Expected value is a certain class and the prediction matches.
- True Negatives(TN): Expected value is to not be a certain class and the prediction matches.
- False Positives(FP): Expected value is a certain class but the prediction doesn't match.
- False Negatives(FN): Expected value to not be a certain class but the prediction doesn't match.

The performance of the classifier models are measured using sklearn.metrics library by the metrics listed below and then used for comparing the models for efficiency.

- Accuracy: Ratio of the number of correct predictions to the total number of predictions. It is defined by the formula

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision: Metric useful when cost of false positives are high. It tells how precise the classifier is in predicting true positives. It is defined as

$$\text{Precision} = \frac{\text{True Positives}}{\text{Total predicted positives}} = \frac{TP}{TP + FP}$$

- Recall: Metric useful when cost of false negatives are very high like predicting the mechanism of injury as 'strain, sprain' when it is actually 'poisoning'. It tells how sensitive the model is to positive predictions.

$$\text{Recall} = \frac{\text{True Positives}}{\text{Total expected positives}} = \frac{TP}{TP + FN}$$

- F1-score: Defined as weighted average or harmonic mean of Precision and Recall with a value ranging from 0 to 1, it is useful when the classes are not balanced where each class is represented by an approximately equal number of records.

$$\text{F1} = \frac{2 * Precision * Recall}{Precision + Recall}$$

Table 5.2. Accuracy

	Naive Bayes	BiLSTM	Distil-BERT	BERT
Accuracy	0.60	0.62	0.37	0.35

The calculated values for precision, recall and f1-score for each class is shown in tables 5.3, 5.4 and 5.5. Based on the analysis, it can be seen that BiLSTM model performs best when compared to the other models. To improve performance of the language models, data was duplicated to create a balanced model and then tested, but it did not yield better results. One of the possible reasons for low accuracy in the language models is that the embeddings used for fine-tuning were trained on documents that contained simple english words unlike the abbreviated medical terminology and phrases seen in the healthcare domain.

Table 5.3. Precision

	Naive Bayes	BiLSTM	Distil-BERT	BERT
41 - INGESTION	0.9	0.83	0.81	0.67
42 - ASPIRATION	0.6	0.42	0	0
46 - BURN, ELECTRICAL	0	0	0	0
47 - BURN, NOT SPEC.	0	0	0	0
48 - BURN, SCALD	0.85	0.83	0.66	0.67
49 - BURN, CHEMICAL	0.77	0.42	0	0
50 - AMPUTATION	0.75	0.84	0	0
51 - BURNS, THERMAL	0.72	0.65	0.64	0.57
52 - CONCUSSION	0.93	0.96	0.41	0.35
53 - CONTUSIONS, ABR.	0.84	0.86	0.56	0.61
54 - CRUSHING	0.55	0.81	0	0
55 - DISLOCATION	0.99	0.9	0.86	0.54
56 - FOREIGN BODY	0.84	0.84	0.61	0.66
57 - FRACTURE	0.8	0.96	0.49	0.4
58 - HEMATOMA	0.91	0.9	0	0
59 - LACERATION	0.81	0.94	0.51	0.27
60 - DENTAL INJURY	0.87	0.89	0	0
61 - NERVE DAMAGE	0	0.91	0	0
62 - INTERNAL INJURY	0.7	0.88	0.48	0.38
63 - PUNCTURE	0.88	0.8	0.88	0.73
64 - STRAIN, SPRAIN	0.85	0.94	0.45	0.47
65 - ANOXIA	0.86	0.89	0.62	0.6
66 - HEMORRHAGE	0	0.84	0	0
67 - ELECTRIC SHOCK	0	0.82	0	0
68 - POISONING	0.87	0.8	0.48	0.43
69 - SUBMERSION	0.88	0.95	0.91	0
71 - OTHER	0.32	0.33	0.25	0.37
72 - AVULSION	0.91	0.86	1	0
73 - RADIATION	0	0.97	0	0
74 - DERMA/CONJUNCT	0.92	0.86	0.83	0.69

Table 5.4. Recall

	Naive Bayes	BiLSTM	Distil-BERT	BERT
41 - INGESTION	0.58	0.3	0.38	0.41
42 - ASPIRATION	0.01	0.09	0	0
46 - BURN, ELECTRICAL	0	0	0	0
47 - BURN, NOT SPEC.	0	0	0	0
48 - BURN, SCALD	0.49	0.22	0.39	0.25
49 - BURN, CHEMICAL	0.16	0.15	0	0
50 - AMPUTATION	0.01	0.36	0	0
51 - BURNS, THERMAL	0.52	0.53	0.35	0.29
52 - CONCUSSION	0.48	0.52	0.13	0.16
53 - CONTUSIONS, ABR.	0.52	0.61	0.19	0.17
54 - CRUSHING	0.03	0.1	0	0
55 - DISLOCATION	0.26	0.35	0.03	0.01
56 - FOREIGN BODY	0.53	0.38	0.34	0.25
57 - FRACTURE	0.6	0.57	0.24	0.31
58 - HEMATOMA	0.09	0.39	0	0
59 - LACERATION	0.63	0.64	0.46	0.74
60 - DENTAL INJURY	0.08	0.21	0	0
61 - NERVE DAMAGE	0	0.28	0	0
62 - INTERNAL INJURY	0.56	0.53	0.3	0.31
63 - PUNCTURE	0.51	0.32	0.18	0.2
64 - STRAIN, SPRAIN	0.61	0.55	0.4	0.33
65 - ANOXIA	0.19	0.21	0.07	0.1
66 - HEMORRHAGE	0	0.19	0	0
67 - ELECTRIC SHOCK	0	0.18	0	0
68 - POISONING	0.37	0.44	0.35	0.29
69 - SUBMERSION	0.2	0.2	0.02	0
71 - OTHER	0.83	0.96	0.74	0.38
72 - AVULSION	0.24	0.26	0	0
73 - RADIATION	0	0.22	0	0
74 - DERMA/CONJUNCT	0.47	0.43	0.22	0.23

Table 5.5. F1-Score

	Naive Bayes	BiLSTM	Distil-BERT	BERT
41 - INGESTION	0.71	0.44	0.51	0.51
42 - ASPIRATION	0.01	0.15	0	0
46 - BURN, ELECTRICAL	0	0	0	0
47 - BURN, NOT SPEC.	0	0	0	0
48 - BURN, SCALD	0.62	0.35	0.49	0.37
49 - BURN, CHEMICAL	0.27	0.22	0	0
50 - AMPUTATION	0.02	0.5	0	0
51 - BURNS, THERMAL	0.6	0.58	0.45	0.38
52 - CONCUSSION	0.63	0.68	0.2	0.22
53 - CONTUSIONS, ABR.	0.65	0.71	0.28	0.27
54 - CRUSHING	0.05	0.18	0	0
55 - DISLOCATION	0.41	0.51	0.05	0.01
56 - FOREIGN BODY	0.65	0.52	0.43	0.36
57 - FRACTURE	0.69	0.71	0.32	0.35
58 - HEMATOMA	0.16	0.55	0	0
59 - LACERATION	0.71	0.76	0.49	0.39
60 - DENTAL INJURY	0.14	0.34	0	0
61 - NERVE DAMAGE	0	0.42	0	0
62 - INTERNAL INJURY	0.62	0.67	0.37	0.34
63 - PUNCTURE	0.64	0.45	0.31	0.32
64 - STRAIN, SPRAIN	0.71	0.7	0.42	0.38
65 - ANOXIA	0.32	0.34	0.13	0.17
66 - HEMORRHAGE	0	0.31	0	0
67 - ELECTRIC SHOCK	0	0.3	0	0
68 - POISONING	0.52	0.56	0.4	0.35
69 - SUBMERSION	0.33	0.33	0.04	0
71 - OTHER	0.47	0.49	0.37	0.37
72 - AVULSION	0.38	0.39	0	0
73 - RADIATION	0	0.36	0	0
74 - DERMA/CONJUNCT	0.62	0.57	0.34	0.35

6. CONCLUSION

With time and accuracy being of highest importance in an emergency situation, this research paper presents a solution to assist the emergency medical service staff in completing their electronic reports automatically and providing them with an option to review and modify incorrect mappings. By custom engineering existing techniques like Metamap and syntactic dependency parser, medical narrative texts are transformed into a 24 concept field and value pair. As part of the evaluation process, interaction with an experienced EMS professional provided more information about the different terminologies used among EMS staff which contained words that are unique to an emergency medical service professional. Since the dataset was limited to trauma calls, to automatically retrieve the mechanism of injury from text, 4 classification models were trained on a much larger dataset that contained medical abbreviations, tested and evaluated for performance.

One of the limitations for this research work include the fact that the dataset was comparitively very small and the information content in it included information was more suited to a dispatch unit that requires just enough information about the patient and location to decide which hospital could provide best treatment and care for the patient which is more conversational and concise compared to a detailed report of the patient provided by the first responder. This was resolved to a certain limit by editing each transcript to remove irrelevant information. Also, since metamap is trained on general healthcare documents, it was unable to capture specific information related to EMS. To resolve that, specific semantic types were chosen and filtered for certain words and then later combined with syntactic dependency parser to provide concept-specific phrases. While analyzing the outputs, it was noticed that the system produced a low score for a few concepts like patient condition, mechanism of injury, etc. One of the reasons for the low score is that such values are populated by human understanding of the medical narrative and was not retrieved as phrases present in text. Since the algorithm discussed in this paper focuses only on extraction, concept values that require human knowledge and understanding to derive a conclusion based on certain other parameters, have a low performance score. With limitation on time, using pre-trained models trained on the english vocabulary for medical abbreviated text caused the efficiency of the language models to be significantly low.

Future improvements to the work done in this research include creating and using a separate lexicon that contains EMS terminologies, increasing the quantity and quality of dataset by retrieving recordings of detailed audio report provided by the field crews either at the incident scene or while reporting to the medical professional at the hospital. Once dataset quality and size is increased and model is trained on EMS ontology, additional constraints on certain fields could be engineered, verified and tested. For example, in an EMR report chief complaint must never be none and in cases when patient says that he/she has no complaint, the system should be able to auto generate a summary of patient problem as the chief complaint. Another improvement is for fields like pupils and lungs, with more data in hand, information regarding pupils and eyes could be split into right and left sides. With more interactions with EMS staff, it is also possible to highlight certain fields that are required for each patient case and provide them with options to choose from and in doing so, minimizing typing effort and errors caused while typing in a high stress environment. The performance of language models can be significantly improved when trained on data containing the EMS terminologies.

Although this algorithms focuses more on the EMS domain, the underlying methodology can be custom engineered to develop automatic concept extraction in other domains that require entity extraction from raw text. Despite the fact that the model has certain limitations, it serves it's intended purpose as an assistant to EMS staff for automated EMR form generation. With information captured at the incident site, the first responder is prone to less errors in analyzing the generated form and modifying it for any corrections since it removed the need to memorize the details of the case.

When combined with a wearable device that records audio and captures images of the incident, converting the voice to text, processing the text to generate concept-specific field value pairs and transforming the generated output to an EMR form, this model can be used in the field.

REFERENCES

[1] DHS, *Emergency medical services*, Last accessed : 05/06/2021. [Online]. Available: https://www.in.gov/dhs/ems/.

[2] P. Medicine, *What is door-to-balloon time?* Last accessed : 31/05/2021, 2017. [Online]. Available: https://www.pennmedicine.org/updates/blogs/heart-and-vascular-blog/2017/january/door-to-balloon.

[3] Toshiba, *Smart glass*, Last accessed : 01/06/2021. [Online]. Available: https://us.dynabook.com/smartglasses/products/index.html.

[4] M. C. Taylor, *Mont. paramedicine program lab uses home, ambulance for training*, Last accessed : 01/06/2021, 2019. [Online]. Available: https://www.emsworld.com/news/1222129/mont-paramedicine-program-lab-uses-home-ambulance-training.

[5] S. Fu, D. Chen, H. He, S. Liu, S. Moon, K. J. Peterson, F. Shen, L. Wang, Y. Wang, A. Wen, Y. Zhao, S. Sohn, and H. Liu, "Clinical concept extraction: A methodology review," *Journal of Biomedical Informatics*, vol. 109, p. 103 526, 2020, Last accessed : 23/04/2021, ISSN: 1532-0464. DOI: https://doi.org/10.1016/j.jbi.2020.103526. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1532046420301544.

[6] K. Yu, G. Guan, and M. Zhou, "Resume information extraction with cascaded hybrid model," in *ACL*, 2005.

[7] S. Mermelstein, P. Street, and Newton, "Information extraction by text classification," 2001.

[8] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, "Dbpedia: A nucleus for a web of open data," in *The Semantic Web*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 722–735, ISBN: 978-3-540-76298-0.

[9] N. Nakashole, M. Theobald, and G. Weikum, "Scalable knowledge harvesting with high precision and high recall," Last accessed : 30/05/2021, Jan. 2011, pp. 227–236. DOI: 10.1145/1935826.1935869.

[10] C. Virmani, A. Pillai, and D. Juneja, "Extracting information from social network using nlp," *International Journal of Computational Intelligence Research*, vol. 13, 2017.

[11] A. Gupta, I. Banerjee, and D. L. Rubin, "Automatic information extraction from unstructured mammography reports using distributed semantics," *Journal of Biomedical Informatics*, vol. 78, pp. 78–86, 2018, Last accessed : 21/03/2021, ISSN: 1532-0464. DOI: https://doi.org/10.1016/j.jbi.2017.12.016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1532046417302903.

[12] S. M. Preum, S. Shu, H. Alemzadeh, and J. A. Stankovic, "Emscontext: Ems protocol-driven concept extraction for cognitive assistance in emergency response," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 08, 2020, Last accessed : 23/04/2021. DOI: 10.1609/aaai.v34i08.7048. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/7048.

[13] ImageTrend, *Imagetrend*, Last accessed : 31/05/2021. [Online]. Available: https://www.imagetrend.com/.

[14] Collectivemedical, *Helping patients get the care they need and reducing inappropriate use of emergency departments*, Last accessed : 25/01/2021. [Online]. Available: https://collectivemedical.com/ed-utilization/.

[15] M. A. Rahman, S. M. Preum, R. Williams, H. Alemzadeh, and J. A. Stankovic, "Grace: Generating summary reports automatically for cognitive assistance in emergency response," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 08, 2020, Last accessed : 13/11/2020. DOI: 10.1609/aaai.v34i08.7049. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/7049.

[16] E. Soysal, J. Wang, M. Jiang, Y. Wu, S. Pakhomov, H. Liu, and H. Xu, "CLAMP – a toolkit for efficiently building customized clinical natural language processing pipelines," *Journal of the American Medical Informatics Association*, vol. 25, no. 3, pp. 331–336, Nov. 2017, Last accessed : 18/04/2021, ISSN: 1527-974X. DOI: 10.1093/jamia/ocx132. eprint: https://academic.oup.com/jamia/article-pdf/25/3/331/34150625/ocx132.pdf. [Online]. Available: https://doi.org/10.1093/jamia/ocx132.

[17] W. Rogers, *Installing and running the public version of metamap*, Last accessed : 31/10/2020, 2020. [Online]. Available: https://metamap.nlm.nih.gov/Docs/README.html.

[18] G. Savova, J. Masanz, P. Ogren, J. Zheng, S. Sohn, K. Kipper-Schuler, and C. Chute, "Mayo clinical text analysis and knowledge extraction system (ctakes): Architecture, component evaluation and applications," *Journal of the American Medical Informatics Association : JAMIA*, vol. 17, pp. 507–13, Sep. 2010, Last accessed : 21/02/2021. DOI: 10.1136/jamia.2009.001560.

[19] GeeksForGeeks, *Naive bayes classifiers*, Last accessed : 31/05/2021, 2020. [Online]. Available: https://www.geeksforgeeks.org/naive-bayes-classifiers/.

[20] Y. Goldberg and O. Levy, *Word2vec explained: Deriving mikolov et al.'s negative-sampling word-embedding method*, Last accessed : 11/05/2021, 2014. [Online]. Available: http://arxiv.org/abs/1402.3722.

[21] Z. Huang, W. Xu, and K. Yu, “Bidirectional LSTM-CRF models for sequence tagging,” *CoRR*, vol. abs/1508.01991, 2015, Last accessed : 09/03/2021. arXiv: 1508.01991. [Online]. Available: http://arxiv.org/abs/1508.01991.

[22] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, *Bert: Pre-training of deep bidirectional transformers for language understanding*, Last accessed : 26/05/2021, 2018. [Online]. Available: http://arxiv.org/abs/1810.04805.

[23] D. Saini, A. Sandhu, M. Gori, and H. Orthner, “A study design for comparing electronic patient care report (epcr) with paper pcr in pre-hospital care,” *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, vol. 2005, p. 1103, Feb. 2005.

[24] E. Division, *Bls aic preparation manual*, Last accessed : 05/03/2021, 2016. [Online]. Available: http://www.bbvrs.net/wp-content/uploads/2016/10/BLS-AIC-Preparation-Manual.pdf.

[25] A. R. Aronson, “Effective mapping of biomedical text to the umls metathesaurus: The metamap program,” 2001.

[26] A. R. Aronson and F.-M. Lang, “An overview of metamap: Historical perspective and recent advances,” *Journal of the American Medical Informatics Association*, vol. 17, 2010, Last accessed : 28/05/2021, ISSN: 1067-5027. DOI: 10.1136/jamia.2009.002733. [Online]. Available: https://doi.org/10.1136/jamia.2009.002733.

[27] E. Chiaramello, F. Pinciroli, A. Bonalumi, A. Caroli, and G. Tognola, “Use of “off-the-shelf” information extraction algorithms in clinical informatics: A feasibility study of metamap annotation of italian medical notes,” *Journal of Biomedical Informatics*, vol. 63, pp. 22–32, 2016, Last accessed : 19/05/2021, ISSN: 1532-0464. DOI: https://doi.org/10.1016/j.jbi.2016.07.017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1532046416300715.

[28] B. A. McCray AT Srinivasan S, “Lexical methods for managing variation in biomedical terminologies,” *Proceedings. Symposium on Computer Applications in Medical Care*, 1994, Last accessed : 16/01/2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2247735/pdf/procascamc00001-0249.pdf.

[29] D. Cutting, J. Kupiec, J. Pedersen, and P. Sibun, “A practical part-of-speech tagger,” in *Proceedings of the Third Conference on Applied Natural Language Processing*, ser. ANLC ’92, Last accessed : 25/02/2021, Trento, Italy: Association for Computational Linguistics, 1992, pp. 133–140. DOI: 10.3115/974499.974523. [Online]. Available: https://doi.org/10.3115/974499.974523.

[30] A. R. Aronson and F.-M. Lang, *Metamap semantic types*, Last accessed : 31/05/2021, 2018. [Online]. Available: https://metamap.nlm.nih.gov/Docs.

[31] M. Honnibal, I. Montani, S. Van Landeghem, and A. Boyd, *spaCy: Industrial-strength Natural Language Processing in Python*, Last accessed : 25/03/2021, 2020. DOI: 10.5281/zenodo.1212303. [Online]. Available: https://doi.org/10.5281/zenodo.1212303.

[32] M. Neumann, D. King, I. Beltagy, and W. Ammar, "Scispacy: Fast and robust models for biomedical natural language processing," *Proceedings of the 18th BioNLP Workshop and Shared Task*, 2019, Last accessed : 03/03/2021. DOI: 10.18653/v1/w19-5034. [Online]. Available: http://dx.doi.org/10.18653/v1/W19-5034.

[33] Allen, *Scispacy*, Last accessed : 31/05/2021, 2016. [Online]. Available: https://allenai.github.io/scispacy/.

[34] J. Li, Y. Sun, R. Johnson, D. Sciaky, C.-H. Wei, R. Leaman, A. P. Davis, C. Mattingly, T. Wiegers, and Z. lu, "Biocreative v cdr task corpus: A resource for chemical disease relation extraction," *Database*, vol. 2016, baw068, May 2016, Last accessed : 27/05/2021. DOI: 10.1093/database/baw068.

[35] U. Research, *Documentation*, Last accessed : 14/03/2021, 2020. [Online]. Available: http://mlreference.com/dependency-tree-spacy.

[36] D. Jurafsky and J. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Dec. 2020, vol. 3.

[37] M. Honnibal and I. Montani, *Spacy 101: Everything you need to know*, Last accessed : 25/03/2021, 2015. [Online]. Available: https://spacy.io/usage/spacy-101.

[38] Manning, C. D, M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky, *The stanford corenlp natural language processing toolkit*, Last accessed : 20/03/2021, 2014. [Online]. Available: https://corenlp.run/.

[39] Guru99, *Pos tagging with nltk and chunking in nlp*, Last accessed : 22/03/2021, 2021. [Online]. Available: https://www.guru99.com/pos-tagging-chunking-nltk.html.

[40] de Marneffe, Marie-Catherine, Manning, and C. D., *Stanford typed dependencies manual*, 2008.

[41] A. Kuchling, *Regular expression howto*, Last accessed : 11/03/2021, 2021. [Online]. Available: https://docs.python.org/3/howto/regex.html#regular-expression-howto.

[42] Regex101, *Regex online tester*, Last accessed : 31/05/2021, 2021. [Online]. Available: https://regex101.com/.

[43] M. Honnibal and I. Montani, *Rue based entity recognition*, Last accessed : 25/03/2021, 2021. [Online]. Available: https://spacy.io/usage/rule-based-matching#entityruler.

[44] A. Kormilitzin, N. Vaci, Q. Liu, and A. J. Nevado-Holgado, "Med7: A transferable clinical natural language processing model for electronic health records," *CoRR*, 2020, Last accessed : 28/11/2020. [Online]. Available: https://arxiv.org/abs/2003.01271.

[45] M. D. Pietro, *Text classification with nlp: Tf-idf vs word2vec vs bert*, Preprocessing, Model Design, Evaluation, Explainability for Bag-of-Words, Word Embedding, Language models, Last accessed : 04/02/2021, 2020. [Online]. Available: https://towardsdatascience.com/text-classification-with-nlp-tf-idf-vs-word2vec-vs-bert-41ff868d1794.

[46] P. W. Code, *Bilstm explained*, Last accessed : 29/04/2021, 2020. [Online]. Available: https://paperswithcode.com/method/bilstm.

[47] GLUON, *Fine-tuning pre-trained bert models*, Last accessed : 23/04/2021, 2018. [Online]. Available: https://nlp.gluon.ai/examples.

[48] K. Navarro, *Blood glucose test for altered mental status*, Last accessed : 24/04/2021, 2013. [Online]. Available: https://www.ems1.com/ems-products/ambulance-disposable-supplies/articles/blood-glucose-test-for-altered-mental-status-CNsaqcanCktCr48m/.

[49] JEMS, *The how, what and why of ems pulse oximetry*, Last accessed : 27/02/2021, 2017. [Online]. Available: https://www.jems.com/patient-care/the-how-what-and-why-of-ems-pulse-oximetry/.

[50] C. for Disease Control and Prevention, *Glasgow coma scale*, Last accessed : 01/03/2021. [Online]. Available: https://www.cdc.gov/masstrauma/resources/gcs.pdf.

[51] M. DC., *Clinical methods: The history, physical, and laboratory examinations. 3rd edition*, Last accessed : 17/05/2021, 1990. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK320/.

[52] E. Pete, *Emt-b acronyms*, Last accessed : 20/04/2021, 2009. [Online]. Available: https://www.cram.com/flashcards/emt-b-acronyms-2770852.

[53] A. Foster, V. Florea, C. Fahrenbruch, J. Blackwood, and T. D. Rea, "Availability and accuracy of ems information about chronic health and medications in cardiac arrest," eng, *The western journal of emergency medicine*, 2017, Last accessed : 22/03/2021, ISSN: 1936-9018. DOI: 10.5811/westjem.2017.5.33198. [Online]. Available: https://doi.org/10.5811/westjem.2017.5.33198.

[54] M. Plus, *Breath sounds*, Last accessed : 19/05/2021. [Online]. Available: https://medlineplus.gov/ency/article/007535.htm.

[55] B. Bachenheimer and K. Dauernheim, *Everyone gets an auscultation*, Last accessed : 21/04/2021. [Online]. Available: https://www.emsworld.com/article/1225424/everyone-gets-auscultation.

[56] M. Colbeck, *The paramedic neurological assessment*, The General Assessment - The eyes: Pupils, Slide 31 of 104, Last accessed : 14/04/2021, 2019. [Online]. Available: https://www.slideshare.net/MarcColbeck/the-paramedic-neurological-assessment.

[57] U. C. P. S. Commission, *National electronic injury surveillance system (neiss)*, Last accessed : 04/04/2021, 2020. [Online]. Available: https://www.cpsc.gov/Research--Statistics/NEISS-Injury-Data.

[58] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.