

Building a Corpus for Sentiment Analysis of Surgical Operation Notes

Jake Albert

Advisors: Dragomir Radev, Wade Schulz



Cardiac Presentation/Symptoms - At Time Of This Admission

Stable Angina	Unstable Angina	Angina equivalent
Non-ST Elevation MI (Non-STEMI)	ST Elevation MI (STEMI)	
Other	No Symptoms	

Cardiac Symptoms - At Time Of Surgery

Stable Angina	Unstable Angina	Angina equivalent
Non-ST Elevation MI (Non-STEMI)	ST Elevation MI (STEMI)	
Other	No Symptoms	

Prior MI

Yes	No	Unknown
-----	----	---------

Cardiac Arrhythmia

Yes	No	Unknown
-----	----	---------

RF-Chronic Lung Disease

Mild	Moderate
Severe	Lung disease documented, severity unknown
No	Unknown

Risk Model and Variables - STS Adult Cardiac Surgery Database Version 2.81

RISK SCORES

[About the STS Risk Calculator](#)

Procedure: CAB Only

Risk of Mortality: 0.264%

Morbidity or Mortality: 3.809%

Long Length of Stay: 1.065%

Short Length of Stay: 81.91%

Permanent Stroke: 0.165%

Prolonged Ventilation: 2.37%

DSW Infection: 0.165%

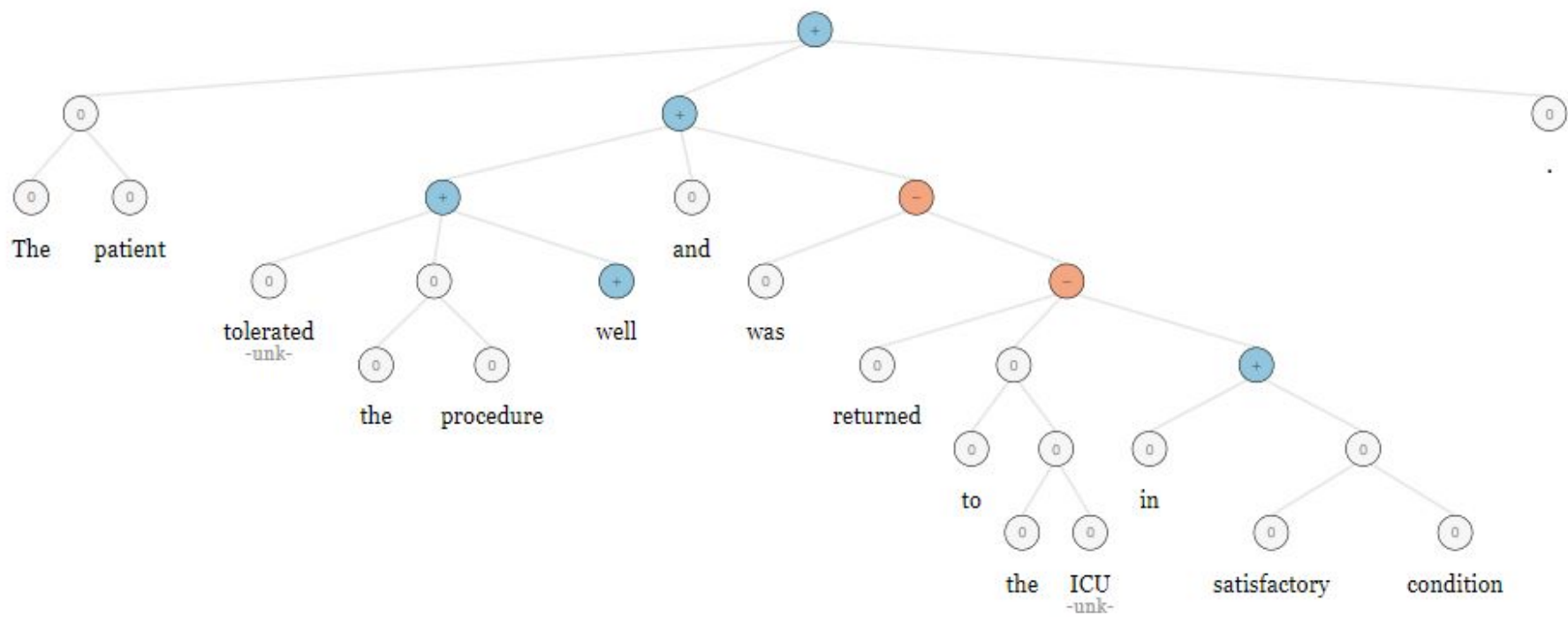
Renal Failure: 0.548%

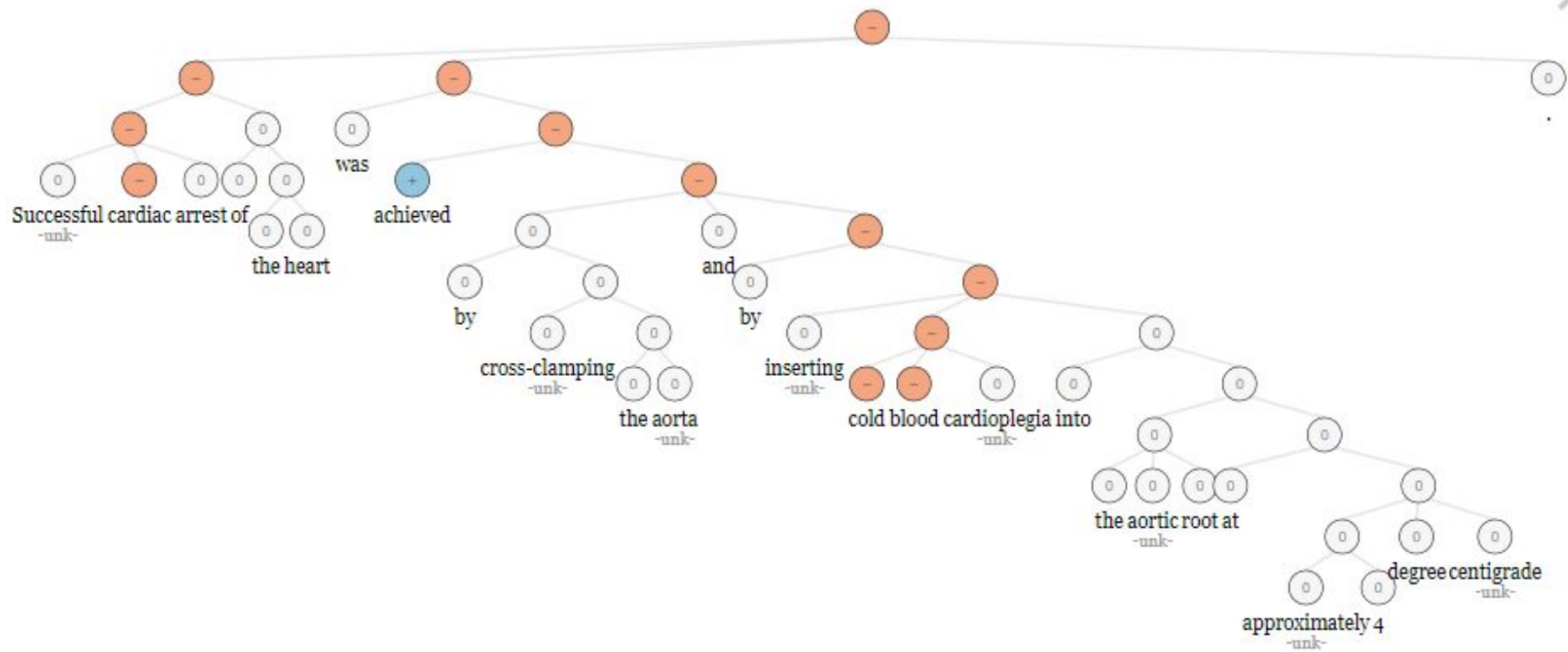
Reoperation: 2.199%

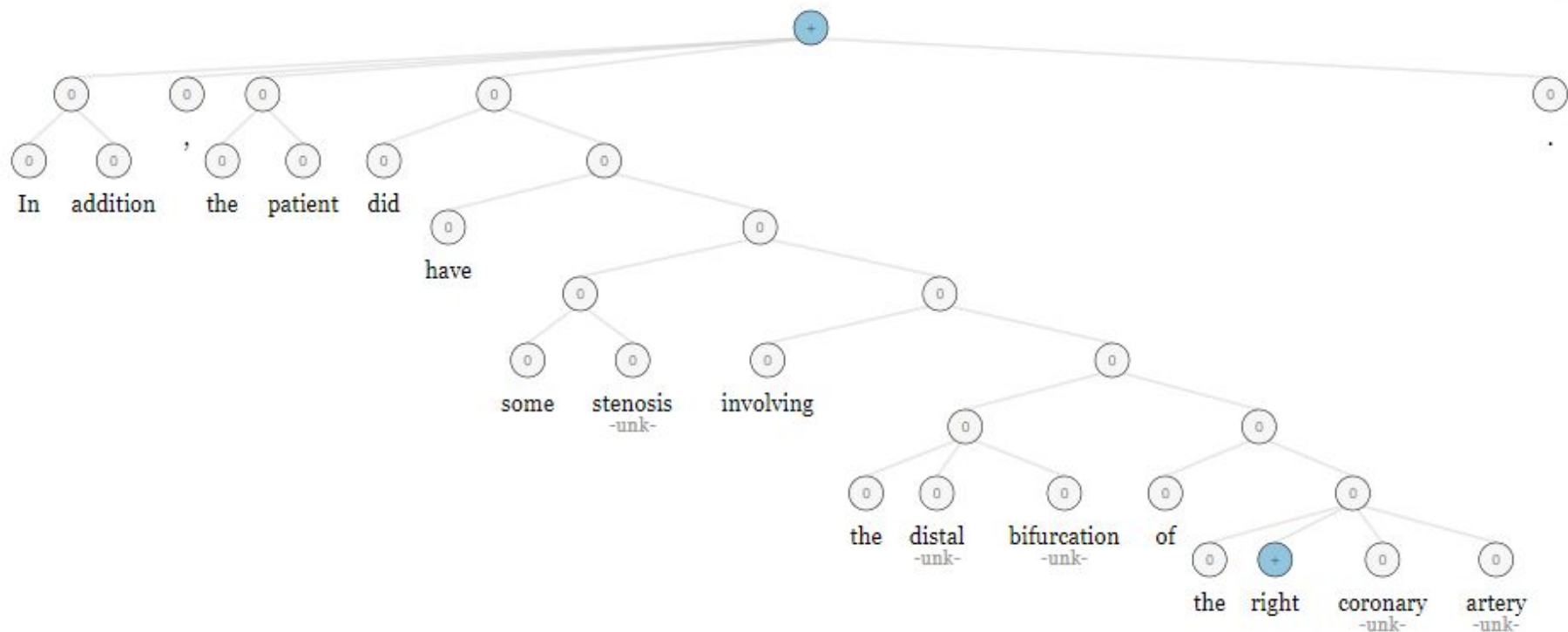
PRINT

CLEAR

First, the PDA was exposed, and the right coronary artery was found to have moderate calcification throughout its course to the terminal bifurcation. Arteriotomy was performed on the PDA, which tolerated placement of a 1.5 French probe. We first constructed an anastomosis of the end of the reversed saphenous vein with 7-0 prolene. Next a 5 mm arteriotomy punch was used to expand a small stab incision in the ascending aorta, and an anastomosis was created between the saphenous vein and the ascending aorta with 7-0 prolene sutures. Next, we exposed the LAD, and arteriotomy of the vessel was performed. An end of the internal mammary was approximated to the arteriotomy and a side-to-side anastomosis was performed using 8-0 running prolene. The internal mammary was tacked to the epicardium using its pedicle. The patient was given a dose of retrograde followed by antegrade cardioplegia, and de-airing was performed. The aortic cross-clamp was removed, and the heart was noted to resume spontaneous coordinated contractions. Temporary pacing wires were placed, followed by placement of Blake drains in the medial section. Two chest tubes were placed in the mediastinum, with two other tubes placed in the bilateral pleural cavities. V-pacing wires were also placed, and the patient was weaned off cardiopulmonary bypass without complication. All hemodynamic monitors, including intraoperative TEE, baseline EKG, and pulmonary monitoring were found to be normal. The pericardium and pericardial fat were loosely approximated, and the sternotomy was closed using steel wires. Next 0 vicryl sutures were used to reapproximate the musculofascial layer, and 4-0 Monocryl with skin glue were used to reapproximate the skin. Dressings were applied, and the patient was transferred to the CTICU on multiple drips of epinephrine, neosynephrine, and vasopressin. In the CTICU the patient was placed on APRV and overall in critical condition. All counts were correct at the end of the procedure. Bypass time was 120 minutes. All counts were correct at the end of the case.







Two Large Issues

Objectivity/Subjectivity

Sentiment Lexicon/Treebank

Plan: A Corpus of Annotated Operation Notes

Two Domain-Specific Sentiment Categories:

- Complicated/Not Complicated
- Concerning/Not Concerning

	COMPLICATED: NO	COMPLICATED: YES
CONCERNING: NO	<p>“All hemodynamic monitors, including intraoperative TEE, baseline EKG, and pulmonary monitoring were found to be normal.”</p>	<p>“First, the PDA was exposed, and the right coronary artery was found to have moderate calcification throughout its course to the terminal bifurcation.”</p>
CONCERNING: YES	<p>“Dressings were applied, and the patient was transferred to the CTICU on multiple drips of epinephrine, neosynephrine, and vasopressin for hypotension.”</p>	<p>“The aortic cross-clamp was removed, however the heart was not noted to resume spontaneous coordinated contractions.”</p>

Welcome to the op-notes annotator.

Note: Choose File opnote2.txt Load

ANESTHESIA: General.

INDICATIONS FOR OPERATION: The patient is a (XX)-year-old female with diagnosis of acute coronary artery syndrome. She has a history of a prior PTCA, diabetes and hypertension. Cardiac catheterization here showed a 90% stenosis of the left main coronary artery. Ejection fraction was 80%. In addition, the patient did have some stenosis involving the distal bifurcation of the right coronary artery.

DESCRIPTION OF OPERATION: The patient was brought to the operative room and placed supine, then induced under suitable general anesthesia. The patient was then prepped and draped in the usual sterile fashion. A sternotomy incision was made and the sternum divided with use of the sternal saw. Using a self-retracting retractor, the left internal mammary artery was mobilized on the pedicle, which included a slip of the endothoracic fascia as well as the accompanied vein. The length of the left greater saphenous vein was harvested endoscopically by first making a small 1 inch transverse incision over the medial aspect of the left knee joint. <<<< The vein was identified and harvested endoscopically. >>>>

Systemic EPO was administered. The pericardium opened in the midline and the distal ascending thoracic aorta cannulated. A two-stage venous cannula was placed in the atrium and a pulmonary vent placed for cardiac decompression. The patient was then placed on

Sentence 19:

The vein was identified and harvested endoscopically.

Decisions:

Now rate all of Sentence 19.
Right now, is the surgery:

Not complicated or Complicated

!Comp

Comp

Not concerning or Concerning

!Conc

Conc

Done scoring sentence Redo Sent.

Updates:

> Finished selecting phrases from Sentence 2
> Added "Acute coronary artery syndrome" as: complicated
> Advanced to Sentence 2
> Rated Sentence 1 as: not concerning
> Rated Sentence 1 as: not complicated
> Finished selecting phrases from Sentence 1
> Beginning on Sentence 1
> Loaded Opnote 1

Training Scheme on Sentences

- Naive Bayes Classifier as baseline result
- RNN (word embeddings learned from corpus)
- RNN (word embeddings from BioNLP)

BioNLP vectors

```
In [47]: import gensim
```

```
In [87]: vectorsPath = os.path.join( ".." , 'embeddings' , "wikipedia-pubmed-and-PMC-w2v.bin")  
  
model = gensim.models.KeyedVectors.load_word2vec_format(vectorsPath, binary=True)
```

```
In [91]: testSentence = ["ventricular", "angina", "sternotomy", "anastomosis", "cardioplegia", "revascularization"]  
vectors = [model[w] for w in testSentence]
```

```
print([v.size for v in vectors])  
print([v[0] for v in vectors])
```

```
[200, 200, 200, 200, 200, 200]  
[0.50101435, -0.48678178, 0.28013107, -0.2825149, 0.06619015, -0.092693366]
```

Important Lesson:

Getting enough labeled data takes TIME

Three Tasks for the Future:

Labeled *words* → create a sentiment lexicon

Labeled *sentences* → improve the classifier

Labeled *notes* → find a function for overall sentiment