

# Clinical Text Analysis Using Interactive Natural Language Processing

Gaurav Trivedi  
Intelligent Systems Program  
University of Pittsburgh  
[trivedigaurav@pitt.edu](mailto:trivedigaurav@pitt.edu)

28th March '15

Introduce name - program  
Project

# Overview

- **Problem**
- **Research Goals**
- **Proposed Solution**
- **Methods and Evaluation**
- **Next Steps and Challenges**

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.  
S\_O\_H  
Report ID Patient ID Principal Date Record Type  
1,LG0CGBrZJw6P LG0CGBrZJw6P 20080327 1003 OP  
E\_O\_H  
[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

#### COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic  
NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD  
SEX: female FELLOW:  
ASSISTANT: \*\*NAME[XXX WWW], RN and \*\*NAME[VVV UUU], RN  
DOB: 05/17/1951 WARD: outpatient  
AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II  
INDICATIONS: 1) change in bowel habits  
MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams  
Benadryl IV  
PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were  
thoroughly explained, informed consent was obtained. Digital rectal exam was  
performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was  
introduced through the anus and advanced to the ascending colon. The quality  
of the prep was good. The instrument was then slowly withdrawn as the colon  
was fully examined. It was not possible to reach the cecum.

1. Electronic Medical Records are organized collection of information about individual patients. They serve as invaluable data sources to the clinicians in their day-to-day work.
2. Why free text? Can't we provide some structure?
  1. Running queries is easier
  2. Keeps the flexibility
3. This makes it harder to extract information from them automatically. (Show example).

le FELLOW:  
\*\*NAME[XXX WWW], RN and \*\*NAME[VVV UUU], RN  
7/1951 WARD: outpatient  
DATE: 03/27/2008  
  
Class II  
S: 1) change in bowel habits  
: 200 micrograms Fentanyl IV, 6 milligrams Versed  
V  
N OF COLON: good  
  
After the risks benefits and alternatives of the  
explained, informed consent was obtained. Digit  
and revealed no abnormalities. The Pentax EC-38  
through the anus and advanced to the ascending c  
n was good. The instrument was then slowly with

1. Electronic Medical Records are organized collection of information about individual patients. They serve as invaluable data sources to the clinicians in their day-to-day work.
2. Why free text? Can't we provide some structure?
  1. Running queries is easier
  2. Keeps the flexibility
3. This makes it harder to extract information from them automatically. (Show example).

# Overview

- **Problem**
- Research Goals
- Proposed Solution
- Methods and Evaluation
- Next Steps and Challenges

# The Problem

- Long history of research on natural language processing methods (NLP) in the clinical domain  
[Chapman et. al., 2011]
- But, analyzing unstructured text is a hard problem
  - Accuracy reported in prior work ranges from **70-90%**!
  - **Incorrect predictions** are possible:
    - Absence of enough training examples.
    - Language features (context, etc.)

# The Problem

- **Domain experts** may be able to fix modeling problems, but they may not be familiar with NLP techniques

# The Problem

- Chapman et al. (2011) highlighted **lack of user-centered development** as one of barriers NLP development in the clinical domain
- Need to focus on development of **generalizable and usable NLP software** for medical applications

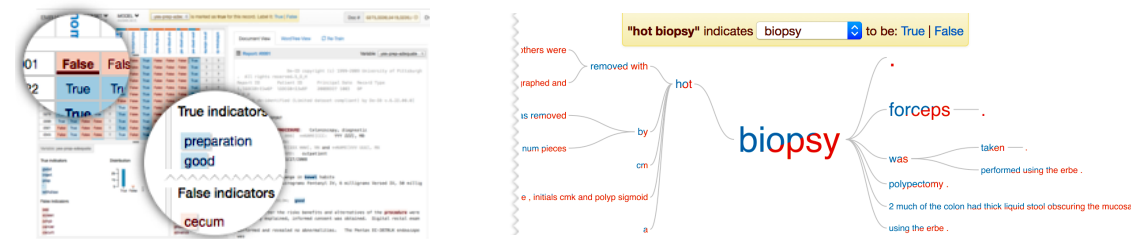
Chapman et al. (2011), *Overcoming barriers to NLP for clinical text*



# Overview

- Problem
- **Research Goals**
- Proposed Solution
- Methods and Evaluation
- Next Steps and Challenges

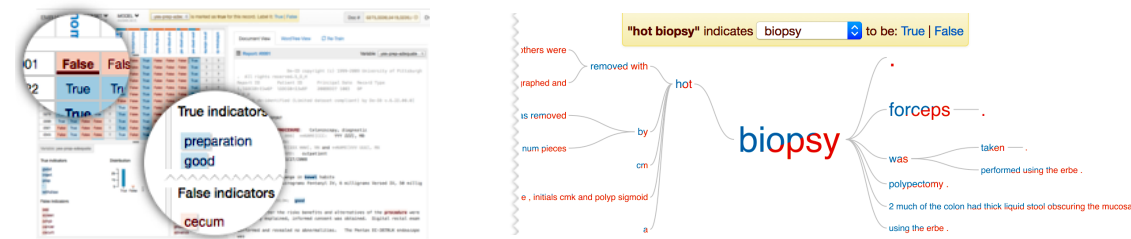
# Research Goals



To explore the design of user interfaces

Let me try to break it down to explain what we are trying to achieve here

# Research Goals

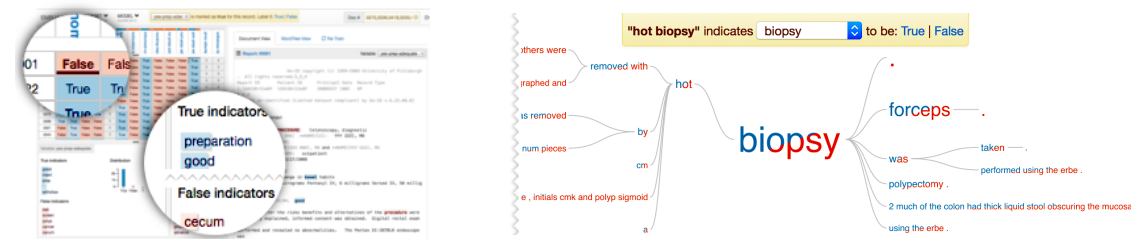


To explore the design of user interfaces:

- *for use by* clinicians and researchers (domain experts)

Let me try to break it down to explain what we are trying to achieve here

# Research Goals

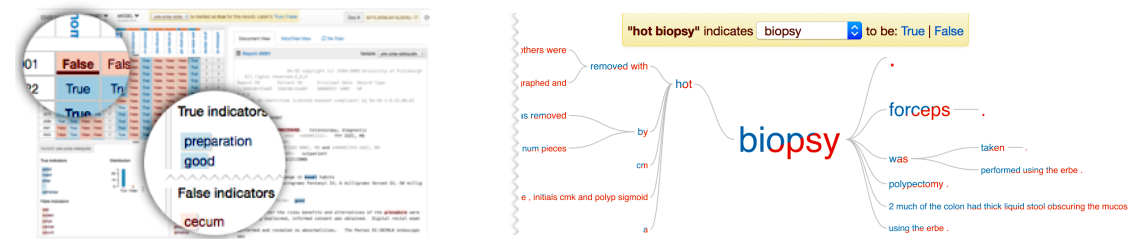


## To explore the design of user interfaces:

- **for use by** clinicians and researchers (domain experts)
- **to support** interactive machine learning on clinical text

Let me try to break it down to explain what we are trying to achieve here

# Research Goals



## To explore the design of user interfaces:

- **for use by** clinicians and researchers (domain experts)
- **to support** interactive machine learning on clinical text
- **in order to** enable rapid convergence on accurate models

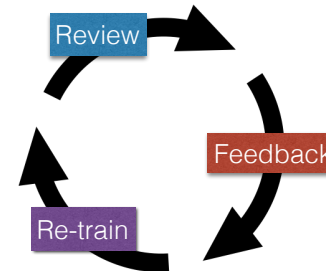
Let me try to break it down to explain what we are trying to achieve here

# Overview

- Problem
- Research Goals
- **Proposed Solution**
- Methods and Evaluation
- Next Steps and Challenges

# Our Solution:

## An Interactive NLP Tool for Clinical Text



- Provide a provision for the end-users to inspect NLP outcomes and make corrections

15

We allow the users to build models iteratively  
Refine them further with every iteration...

# Related Work

## Interactive Machine Learning

- Interactive systems that can learn from their end-users are becoming widespread

Interactive image segmentation [Fails and Osen, 2003], Music Composition [Fiebrink et. al., 2011], Bug Triage [Amershi et. al. 2011], Document Retrieval [Heimerl et. al., 2012] and many more...

Amershi et. al., Power to the people - the role of humans in interactive machine learning (2014)



# Related Work

## Interactive Machine Learning

- Useful when we don't have large amount of training data available for use
- Reviewing and annotating for clinical domain is difficult, requires expertise and time consuming
- We need techniques to obtain good quality training data efficiently

# Related Work

## Visualization and Sensemaking

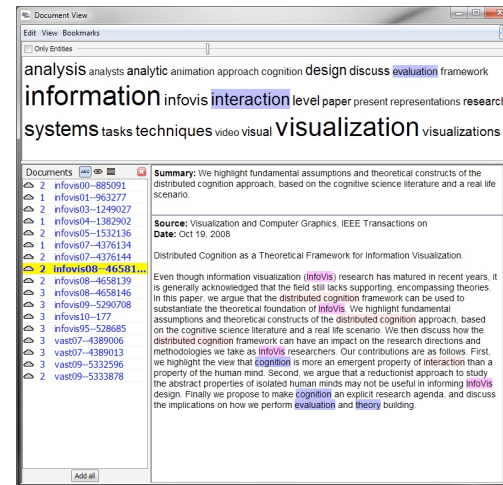
- Visual summary of large amount of text data

This brings to the other related work that can help us with these problems

# Related Work

## Visualization and Sensemaking

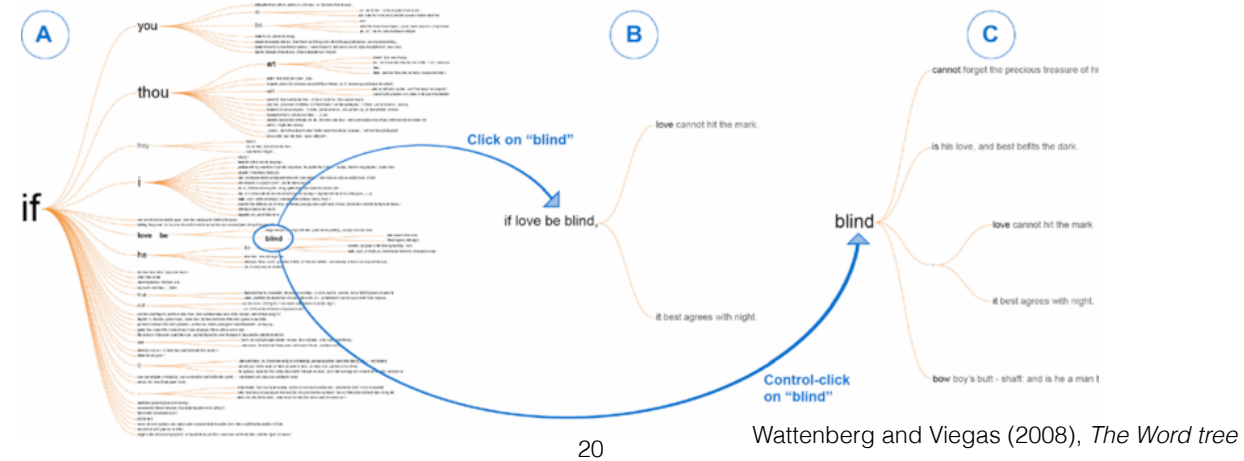
- Visual summary of large amount of text data
- Detailed document level views



# Related Work

## Visualization and Sensemaking

- Visual summary of large amount of text data
- Detailed document level views
- And data-set level views



# Overview

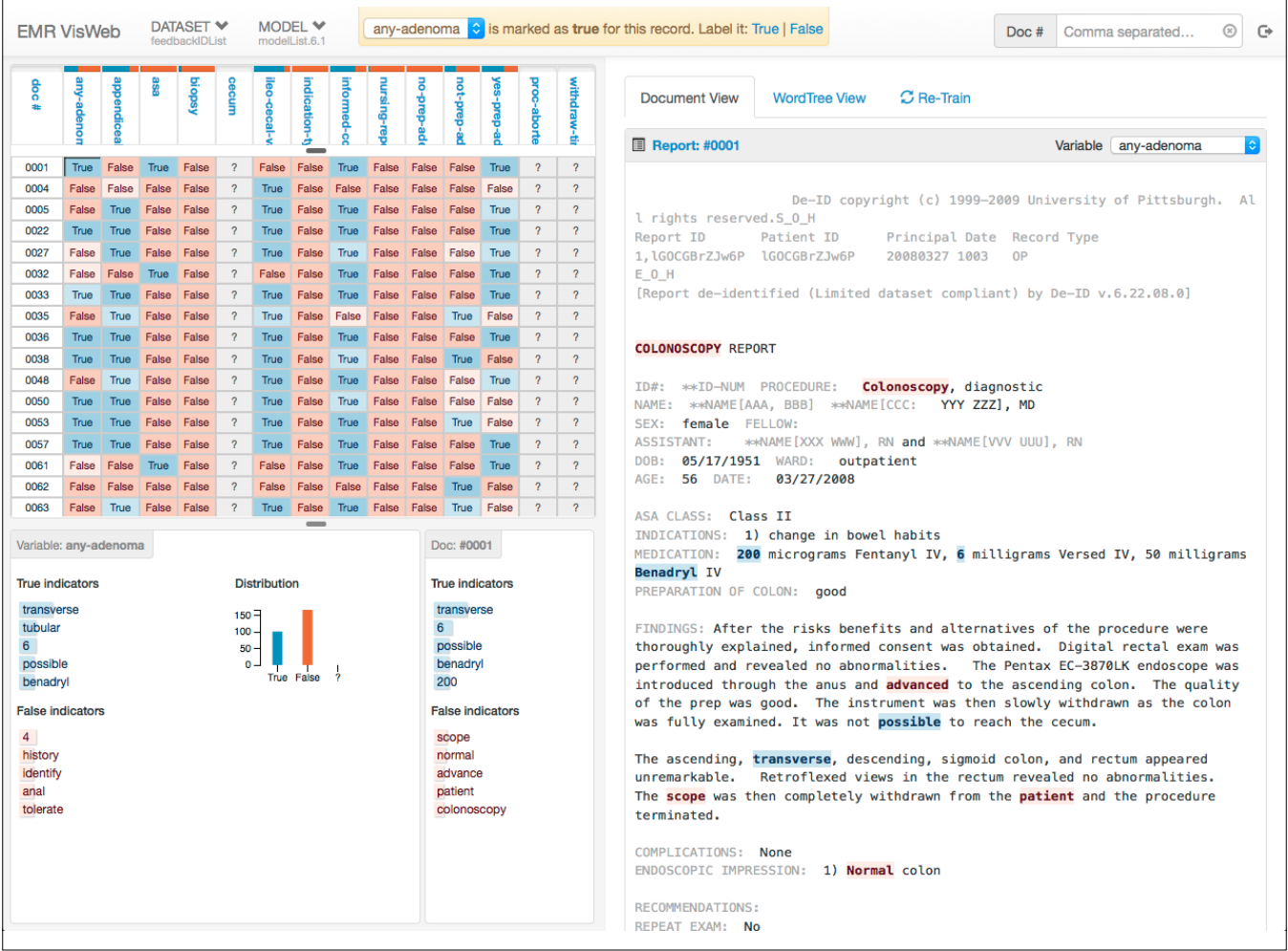
- Problem
- Research Goals
- Proposed Solution
- **Methods and Evaluation**
- Next Steps and Challenges

# Design Requirements

**R1:** The tool should make it easier for machine learning non-experts to work with NLP models.

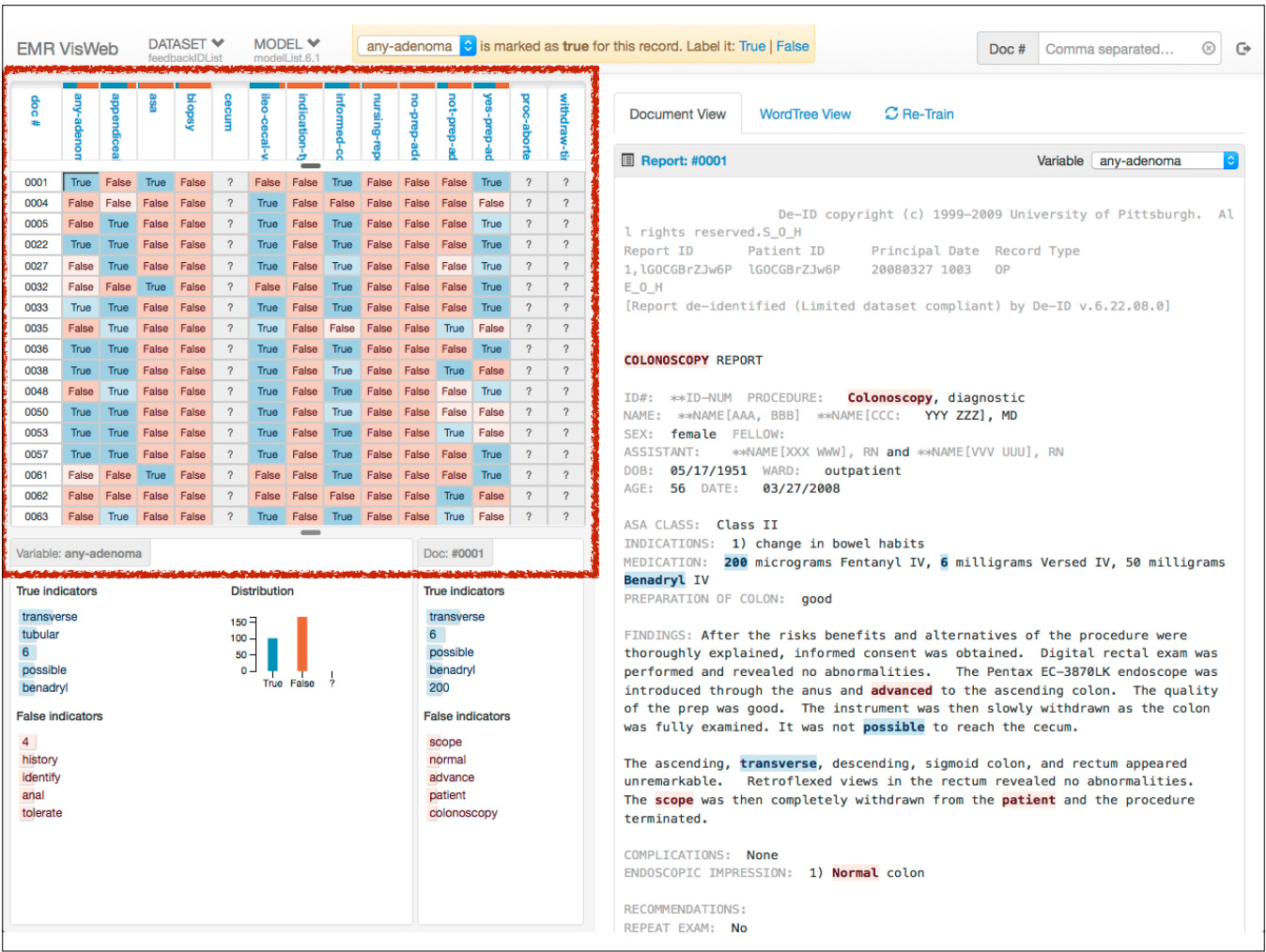
**R2:** It should incorporate efficient mechanisms for annotation and labeling, and also encourage feedback that is consistent and informative.

**R3:** The interactive components should support the entire interactive machine learning loop - i.e. a *review*, *feedback* and *retrain* cycle.



We have developed an interactive web-based tool that facilitates review of binary variables extracted from clinical text.

Here’s a screenshot of it.



The **Grid view** shows the extracted



EMR VisWeb

DATASET  
modelList

MODEL  
modelList 6.1

any adenoma is marked as true for this record. Label it: True | False

Doc # Comma separated...

doc #	any-adenoma	appendiceal	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-cc	nursing-rep	no-prep-adt	no-prep-ad	yes-prep-ad	proc-aborts	withdrawn
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	True	True	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	True	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: any-adenoma

Doc: #0001

True indicators

transverse

tubular

6

possible

benadryl

False indicators

4

history

identify

anal

tolerate

Distribution

True False ?

True indicators

transverse

6

possible

benadryl

200

False indicators

scope

normal

advance

patient

colonoscopy

Document View

WordTree View

Re-Train

Report: #0001

Variable any-adenoma

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_O\_H

Report ID Patient ID Principal Date Record Type

1,lgOCGBrZJw6P lgOCGBrZJw6P 20080327 1003 OP

E\_O\_H

[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic

NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD

SEX: female FELLOW:

ASSISTANT: \*\*NAME[XXX WW], RN and \*\*NAME[VVV UUU], RN

DOB: 05/17/1951 WARD: outpatient

AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II

INDICATIONS: 1) change in bowel habits

MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV

PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined. It was not possible to reach the cecum.

The ascending, transverse, descending, sigmoid colon, and rectum appeared unremarkable. Retroflexed views in the rectum revealed no abnormalities. The scope was then completely withdrawn from the patient and the procedure terminated.

COMPLICATIONS: None

ENDOSCOPIC IMPRESSION: 1) Normal colon

RECOMMENDATIONS:

REPEAT EXAM: No

boolean variables in columns - Example: "Informed Consent" - Whether informed consent was documented in this record or not?

EMR VisWeb

DATASET  
feedbackIDList

MODEL  
modelList.6.1

any-adenoma is marked as true for this record. Label it: True | False

Doc # Comma separated...

doc #	any-adenoma	appendiceal	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-cc	nursing-rep	no-prep-adt	no-prep-ad	yes-prep-ad	proc-aborte	withdraw-tt
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	True	False	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	False	True	?	?

Variable: any-adenoma

Doc: #0001

True indicators

transverse

tubular

6

possible

benadryl

False indicators

4

history

identify

anal

tolerate

Distribution

True indicators

transverse

6

possible

benadryl

200

False indicators

scope

normal

advance

patient

colonoscopy

Document View

WordTree View

Re-Train

Report: #0001

Variable any-adenoma

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_O\_H

Report ID Patient ID Principal Date Record Type

1,lGOCGBrZJw6P lGOCGBrZJw6P 20080327 1003 OP

E\_O\_H

[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic

NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD

SEX: female FELLOW:

ASSISTANT: \*\*NAME[XXX WW], RN and \*\*NAME[VVV UUU], RN

DOB: 05/17/1951 WARD: outpatient

AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II

INDICATIONS: 1) change in bowel habits

MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV

PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined. It was not possible to reach the cecum.

The ascending, transverse, descending, sigmoid colon, and rectum appeared unremarkable. Retroflexed views in the rectum revealed no abnormalities. The scope was then completely withdrawn from the patient and the procedure terminated.

COMPLICATIONS: None

ENDOSCOPIC IMPRESSION: 1) Normal colon

RECOMMENDATIONS:

REPEAT EXAM: No

and individual documents in rows

EMR VisWeb

DATASET  
feedbackIDList

MODEL  
modelList.6.1

any-adenoma is marked as true for this record. Label it: True | False

Doc # Comma separated...

doc #	any-adenoma	appendiceal	asa	biopsy	cecum	ileo-cecal-v	indications	informed-cc	no-prep-ad	no-prep-ad	yes-prep-ad	proc-aborts	withdraw-tt
0001	True	False	True	False	?	False	False	True	False	False	False	True	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?
0032	False	False	True	False	?	False	True	False	False	False	False	True	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?
0035	False	True	False	False	?	True	False	False	False	False	False	True	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?
0048	False	True	False	False	?	True	False	True	False	False	True	True	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?
0061	False	False	True	False	?	False	False	True	False	False	True	True	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?
0063	False	True	False	False	?	True	False	True	False	True	False	False	?

Variable: any-adenoma

Doc: #0001

True indicators

transverse

tubular

6

possible

benadryl

False indicators

4

history

identify

anal

tolerate

Distribution

True indicators

transverse

6

possible

benadryl

200

False indicators

scope

normal

advance

patient

colonoscopy

Document View

WordTree View

Re-Train

Report: #0001

Variable any-adenoma

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_O\_H

Report ID Patient ID Principal Date Record Type

1,lgOCGBrZJw6P lgOCGBrZJw6P 20080327 1003 OP

E\_O\_H

[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic

NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD

SEX: female FELLOW:

ASSISTANT: \*\*NAME[XXX WW], RN and \*\*NAME[VVV UUU], RN

DOB: 05/17/1951 WARD: outpatient

AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II

INDICATIONS: 1) change in bowel habits

MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV

PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined. It was not possible to reach the cecum.

The ascending, transverse, descending, sigmoid colon, and rectum appeared unremarkable. Retroflexed views in the rectum revealed no abnormalities. The scope was then completely withdrawn from the patient and the procedure terminated.

COMPLICATIONS: None

ENDOSCOPIC IMPRESSION: 1) Normal colon

RECOMMENDATIONS:

REPEAT EXAM: No

So this cells shows a prediction value of false against the variable for the record with id 0035.

EMR VisWeb

DATASET  
feedbackIDList

MODEL  
modelList.6.1

any-adenoma is marked as true for this record. Label it: True | False

Doc #

Comma separated...

doc #	any-adenoma	appendiceal	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-cc	nursing-rep	no-prep-adt	not-prep-ad	yes-prep-ad	proc-aborte	withdraw-tt
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	True	False	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: any-adenoma

True indicators

transverse

tubular

6

possible

benadryl

False indicators

4

history

identify

anal

tolerate

Distribution

150

100

50

0

True

False

?

Doc: #0001

True indicators

transverse

6

possible

benadryl

200

False indicators

scope

normal

advance

patient

colonoscopy

Document View

WordTree View

Re-Train

Report: #0001

Variable any-adenoma

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_O\_H

Report ID Patient ID Principal Date Record Type

1,lGOCGBrZJw6P lGOCGBrZJw6P 20080327 1003 OP

E\_O\_H

[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic

NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD

SEX: female FELLOW:

ASSISTANT: \*\*NAME[XXX WW], RN and \*\*NAME[VVV UUU], RN

DOB: 05/17/1951 WARD: outpatient

AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II

INDICATIONS: 1) change in bowel habits

MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV

PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined. It was not possible to reach the cecum.

The ascending, transverse, descending, sigmoid colon, and rectum appeared unremarkable. Retroflexed views in the rectum revealed no abnormalities. The scope was then completely withdrawn from the patient and the procedure terminated.

COMPLICATIONS: None

ENDOSCOPIC IMPRESSION: 1) Normal colon

RECOMMENDATIONS:

REPEAT EXAM: No

Below the grid, we have statistics about the active variable with the distribution of the classifications

EMR VisWeb

DATASET  
feedbackDist

MODEL  
modelList.6.1

any-adenoma is marked as true for this record. Label it: True | False

Doc #

Comma separated...

doc #	any-adenoma	appendiceal	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-cc	nursing-rep	no-prep-adt	no-prep-ad	yes-prep-ad	proc-aborte	withdraw-tt
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	True	False	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	True	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: any-adenoma

True indicators

transverse

tubular

6

possible

benadryl

False indicators

4

history

identify

anal

tolerate

Distribution

True False ?

Doc: #0001

True indicators

transverse

6

possible

benadryl

200

False indicators

scope

normal

advance

patient

colonoscopy

Document View

WordTree View

Re-Train

Report: #0001

Variable any-adenoma

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_O\_H

Report ID Patient ID Principal Date Record Type

1,lGOCGBrZJw6P lGOCGBrZJw6P 20080327 1003 OP

E\_O\_H

[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic

NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD

SEX: female FELLOW:

ASSISTANT: \*\*NAME[XXX WW], RN and \*\*NAME[VVV UUU], RN

DOB: 05/17/1951 WARD: outpatient

AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II

INDICATIONS: 1) change in bowel habits

MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV

PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined. It was not possible to reach the cecum.

The ascending, transverse, descending, sigmoid colon, and rectum appeared unremarkable. Retroflexed views in the rectum revealed no abnormalities. The scope was then completely withdrawn from the patient and the procedure terminated.

COMPLICATIONS: None

ENDOSCOPIC IMPRESSION: 1) Normal colon

RECOMMENDATIONS:

REPEAT EXAM: No

The **Document view** shows the full-text of the patient reports

EMR VisWeb

DATASET  
feedbackIDList

MODEL  
modelList.6.1

any-adenoma is marked as true for this record. Label it: True | False

Doc #

Comma separated...

doc #	any-adenoma	appendiceal	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-cc	nursing-rep	no-prep-adt	not-prep-ad	yes-prep-ad	proc-aborte	withdraw-tt
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	True	False	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: any-adenoma

True indicators

transverse

tubular

6

possible

benadryl

False indicators

4

history

identify

anal

tolerate

Distribution

Doc: #0001

True indicators

transverse

6

possible

benadryl

200

False indicators

scope

normal

advance

patient

colonoscopy

Document View

WordTree View

Re-Train

Report: #0001

Variable: any-adenoma

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_O\_H  
Report ID Patient ID Principal Date Record Type  
1,lGOCGBrZJw6P lGOCGBrZJw6P 20080327 1003 OP  
E\_O\_H  
[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic  
NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD  
SEX: female FELLOW:  
ASSISTANT: \*\*NAME[XXX WW], RN and \*\*NAME[VVV UUU], RN  
DOB: 05/17/1951 WARD: outpatient  
AGE: 56 DATE: 03/27/2008ASA CLASS: Class II  
INDICATIONS: 1) change in bowel habits  
MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV  
PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined. It was not possible to reach the cecum.

The ascending, transverse, descending, sigmoid colon, and rectum appeared remarkable. Retroflexed views in the rectum revealed no abnormalities. The scope was then completely withdrawn from the patient and the procedure terminated.

COMPLICATIONS: None  
ENDOSCOPIC IMPRESSION: 1) Normal colonRECOMMENDATIONS:  
REPEAT EXAM: No

with the indicator terms highlighted;  
Same color scheme as in the grid



EMR VisWeb

DATASET feedbackIDList

MODEL modelList.6.1

yes-prep-adeq

 is marked as true for this record. Label it: True | False

Doc #

Comma separated...

doc #	any-adenom	appendicea	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-cc	nursing-rep	no-prep-adt	no-prep-ad	yes-prep-ad	proc-aborte	withdraw-iti
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	True	False	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	True	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: yes-prep-adequate

True indicators

preparation  
good  
insert  
adequate  
bowel

False indicators

see  
internal  
screen  
significant  
hemorrhoid

Distribution

True

False

?

Doc: #0001

True indicators

preparation  
good  
bowel  
prep  
withdraw

False indicators

cecum  
introduce  
descend  
procedure  
advance

Document View

WordTree View

Re-Train

Report: #0001

Variable: yes-prep-adequate

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_0\_H  
Report ID Patient ID Principal Date Record Type  
1,lG0CGBrZJw6P lG0CGBrZJw6P 20080327 1003 OP  
E\_0\_H  
[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM PROCEDURE: Colonoscopy, diagnostic  
NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD  
SEX: female FELLOW:  
ASSISTANT: \*\*NAME[XXX WWW], RN and \*\*NAME[VVV UUU], RN  
DOB: 05/17/1951 WARD: outpatient  
AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II  
INDICATIONS: 1) change in bowel habits  
MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV  
PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined. It was not possible to reach the cecum.

The ascending, transverse, descending, sigmoid colon, and rectum appeared unremarkable. Retroflexed views in the rectum revealed no abnormalities. The scope was then completely withdrawn from the patient and the procedure terminated.


COMPLICATIONS: None  
ENDOSCOPIC IMPRESSION: 1) Normal colon


RECOMMENDATIONS:  
RPPFAT EXAM: No

Feedback (Type I) can be sent using the yellow control bar on the top,

In some ways this works similar to you marking a mail as spam or not spam in your mailbox.

EMR VisWeb

DATASET  feedbackIDList

MODEL  modelList.6.1


"prep was good" indicates

yes-prep-adeq

to be: True | False

Doc #

Comma separated...



doc #	any-adenom	appendicea	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-cc	nursing-rep	no-prep-ade	no-prep-ade	yes-prep-ade	proc-aborte	withdraw-tt
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	True	False	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: yes-prep-adequate

True Indicators

preparation

good

insert

adequate

bowel

False Indicators

see

internal

screen

significant

hemorrhoid

Distribution

150

100

50

0

True

False

?

Doc: #0001

True Indicators

preparation

good

bowel

prep

withdraw

False Indicators

cecum

introduce

descend

procedure

advance

Document View

WordTree View

Re-Train

Report: #0001

Variable yes-prep-adequate

De-ID copyright (c) 1999-2009 University of Pittsburgh. All rights reserved.S\_O\_H

Report ID Patient ID Principal Date Record Type

1,lG0CGBrZJw6P lG0CGBrZJw6P 20080327 1003 OP

E\_O\_H

[Report de-identified (Limited dataset compliant) by De-ID v.6.22.08.0]

COLONOSCOPY REPORT

ID#: \*\*ID-NUM

PROCEDURE: Colonoscopy, diagnostic

NAME: \*\*NAME[AAA, BBB] \*\*NAME[CCC: YYY ZZZ], MD

SEX: female FELLOW:

ASSISTANT: \*\*NAME[XXX WWW], RN and \*\*NAME[VVV UUU], RN

DOB: 05/17/1951 WARD: outpatient

AGE: 56 DATE: 03/27/2008

ASA CLASS: Class II

INDICATIONS: 1) change in bowel habits

MEDICATION: 200 micrograms Fentanyl IV, 6 milligrams Versed IV, 50 milligrams Benadryl IV

PREPARATION OF COLON: good

FINDINGS: After the risks benefits and alternatives of the procedure were thoroughly explained, informed consent was obtained. Digital rectal exam was performed and revealed no abnormalities. The Pentax EC-3870LK endoscope was introduced through the anus and advanced to the ascending colon. The quality of the prep was good. The instrument was then slowly withdrawn as the colon was fully examined.

"prep was good" indicates "yes-prep-adequate" to be:

True

False

Search using wordtree

COMPLICATIONS: None

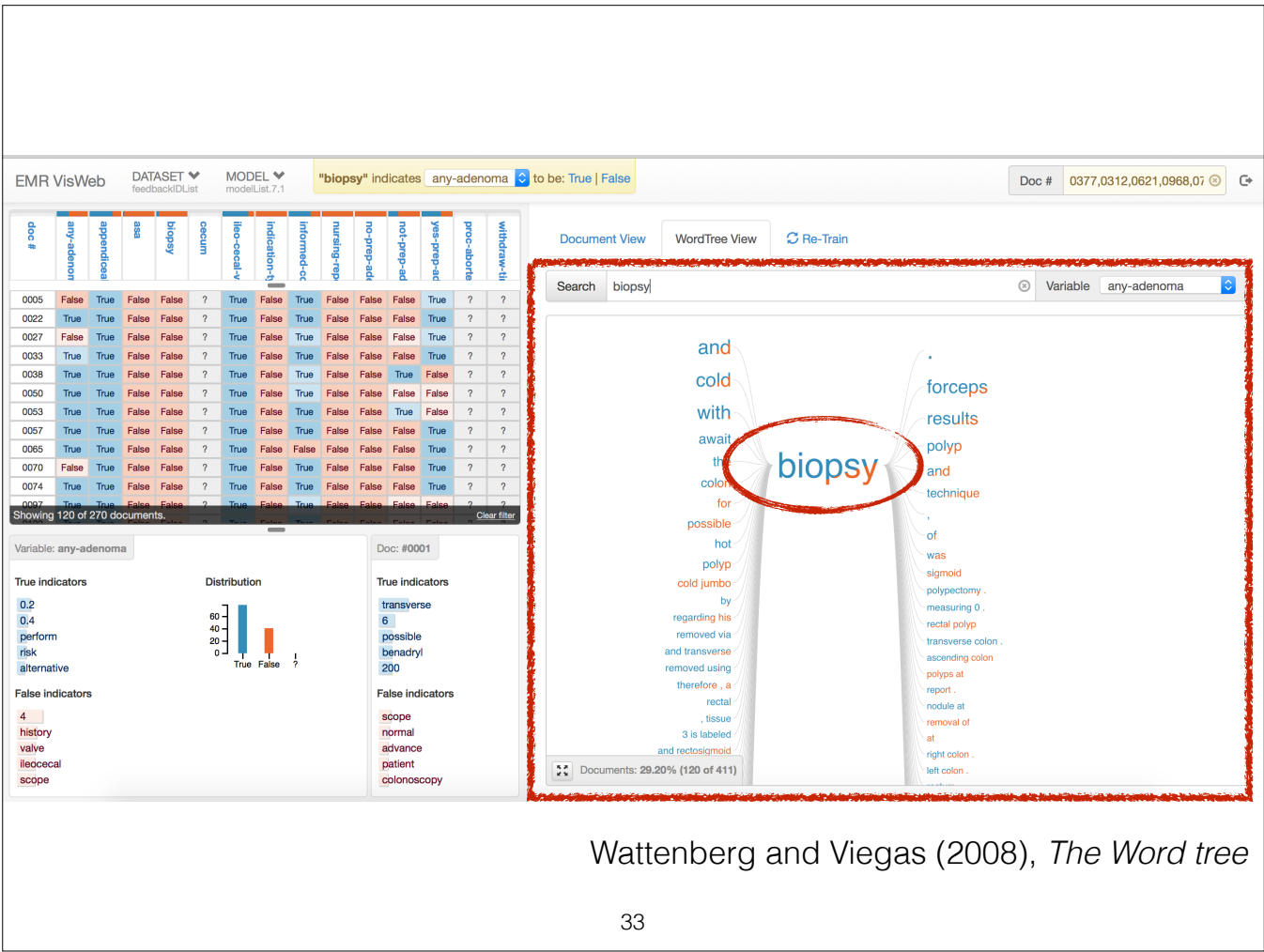
ENDOSCOPIC IMPRESSION:

RECOMMENDATIONS:

REPEAT EXAM: No

Feedback - Highlight span





The **WordTree view** is a type of a text visualization technique that provides the ability to search for and explore word sequence patterns found across the documents in the corpus, and to provide feedback that will be used to retrain NLP models.

EMR VisWeb

DATASET

feedbackIDList

MODEL

modelList.7.1

"cold biopsy" indicates

any-adenoma

to be: True | False

Doc #

0377,0312,0621,0968,07

Document View

WordTree View

Re-Train

Search

biopsy

Variable

any-adenoma

True indicators

0.2

0.4

perform

risk

alternative

False indicators

4

history

valve

ileocecal

scope

Doc: #0001

True indicators

transverse

6

possible

benadryl

200

False indicators

scope

normal

advance

patient

colonoscopy

was

removed

polyp

and

were

colon

also

completely

cm

one

polyps

mm

its

small

and

removed

with

cold

biopsy

forceps

and

polypectomy

technique

Wattenberg and Viegas (2008), *The Word tree*

The **WordTree view** is a type of a text visualization technique that provides the ability to search for and explore word sequence patterns found across the documents in the corpus, and to provide feedback that will be used to retrain NLP models.

EMR VisWeb

DATASET  
feedbackIDList

MODEL  
modelList.6.1

Doc #Comma separated...

doc #	any-admon	appendicea	asa	biopsy	cecum	ileo-cecal-v	indication-q	informed-co	nursing-rep	no-prep-adt	not-prep-adt	yes-prep-adt	proc-abort	withdraw-ti
0001	True	False	True	False	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	False	True	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: yes-prep-adequate

True indicators

preparation  
good  
insert  
adequate  
bowel

False indicators

see  
internal  
screen  
significant  
hemorrhoid

Distribution

150  
100  
50  
0

TrueFalse?

Doc: #0001

True indicators

preparation  
good  
bowel  
prep  
withdraw

False indicators

cecum  
introduce  
descend  
procedure  
advance

Document ViewWordTree ViewRe-Train 2

Feedback List

Mark all documents as unread | Remove bolded change markers

1. #0001 should have **biopsy** marked as **true**.

2. "*prep was good*" indicates **yes-prep-adequate** to be **true**.

Clear All

Re-Train

Training Results

Re-training models. Please wait.

The **Re-Train view** lists user-provided feedback,

EMR VisWeb

DATASET  
feedbackIDList

MODEL  
modelList.7.1

Doc #Comma separated...

doc #	any-adenom	appendicea	asa	biopsy	cecum	ileo-cecal-v	indications-g	informed-cc	nursing-rep	no-prep-adt	not-prep-ad	yes-prep-ad	proc-aborts	withdraw-tt
0001	True	False	True	True	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	False	True	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: any-adenoma

True indicators

transverse  
tubular  
6  
possible  
benadryl

False indicators

4  
history  
identify  
anal  
tolerate

Distribution

150  
100  
50  
0

TrueFalse?

Doc: #0001

True indicators

transverse  
6  
possible  
benadryl  
200

False indicators

scope  
normal  
advance  
patient  
colonoscopy

Document ViewWordTree ViewRe-Train

Feedback List

1. #0001 should have any-adenoma marked as true.

2. "Patient ID" indicates any-adenoma to be true.

3. "Patient ID Pr" indicates any-adenoma to be true.

4. "biopsy" indicates any-adenoma to be true.

5. #0001 should have any-adenoma marked as false.

Clear AllRe-Train

Training Results

✖ You have conflicting labels in your current set of feedback:

Cannot set any-adenoma to be both true and false in Doc #0001!

Cannot set any-adenoma to be both true and false (using "Patient ID") in Doc #0001!

Cannot set any-adenoma to be both true and false (using "Patient ID Pr") in Doc #0001!

Try re-training after making the feedback list consistent. Use the x button next to a feedback to remove it.

including any potential inconsistencies...

EMR VisWeb

DATASET  
feedbackIDList

MODEL  
modelList.7.1

Doc #Comma separated...

doc #	any-adenom	appendicea	asa	biopsy	cecum	ileo-cecal-v	indications-g	informed-cc	nursing-rep	no-prep-adt	not-prep-ade	yes-prep-ade	proc-aborts	withdraw-tt
0001	True	False	True	True	?	False	False	True	False	False	False	True	?	?
0004	False	False	False	False	?	True	False	False	False	False	False	False	?	?
0005	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0022	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0027	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0032	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0033	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0035	False	True	False	False	?	True	False	False	False	False	True	False	?	?
0036	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0038	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0048	False	True	False	False	?	True	False	True	False	False	False	True	?	?
0050	True	True	False	False	?	True	False	True	False	False	False	False	?	?
0053	True	True	False	False	?	True	False	True	False	False	True	False	?	?
0057	True	True	False	False	?	True	False	True	False	False	False	True	?	?
0061	False	False	True	False	?	False	False	True	False	False	False	True	?	?
0062	False	False	False	False	?	False	False	False	False	False	True	False	?	?
0063	False	True	False	False	?	True	False	True	False	False	True	False	?	?

Variable: any-adenoma

True indicators

transverse  
tubular  
6  
possible  
benadryl

False indicators

4  
history  
identify  
anal  
tolerate

Distribution

TrueFalse?

Doc: #0001

True indicators

transverse  
6  
possible  
benadryl  
200

False indicators

scope  
normal  
advance  
patient  
colonoscopy

Document ViewWordTree ViewRe-Train

Feedback List

Mark all documents as unread | Remove bolded change markers

Feedback list is empty.

Training Results

Retraining successfull Now using modelList.7.1

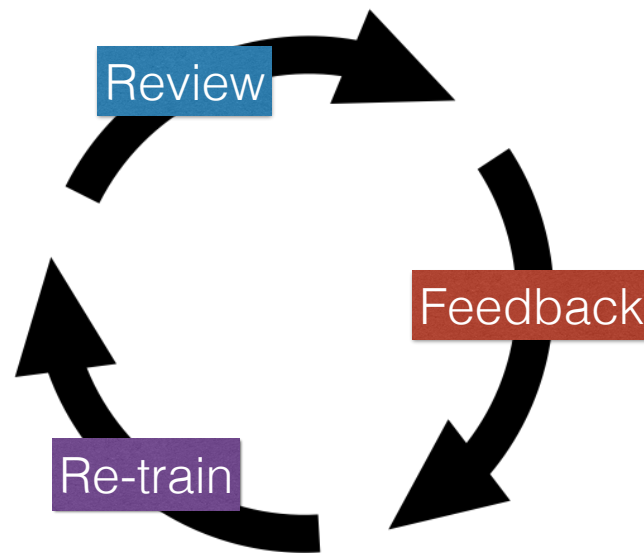
It includes the following feedback:

1. #0001 should have **biopsy** marked as **true**.

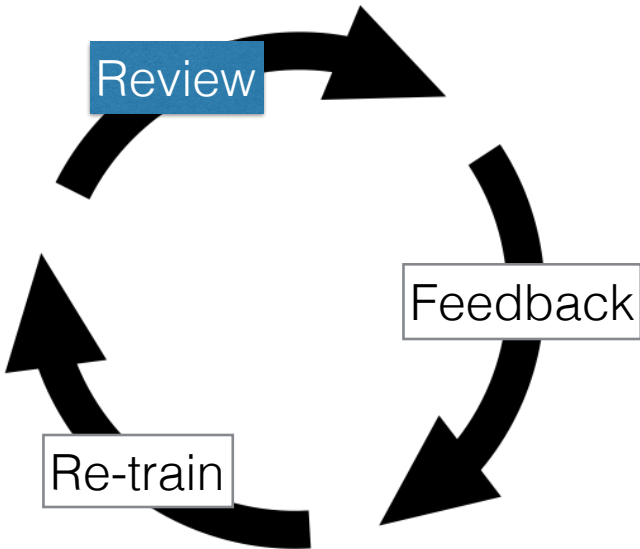
2. "prep was good" indicates **yes-prep-adequate** to be **true**.

and specifies changes in variable assignments due to retraining.

# Feedback Loop



# Feedback Loop



doc #	any adenoma	appendiceal	anal	sigmoid	cecum
0001	False	False	True	False	?
0004	False	False	False	False	?
0005	False	True	False	False	?

the risks benefits and alternatives of the procedure were discussed, informed consent was obtained. Digital rectal exam was normal, no abnormalities. The Pentax EC-3870LK endoscope was inserted to the anus and advanced to the ascending colon. The quality of the exam was good. The instrument was then slowly withdrawn as the colon was completely examined. It was not possible to reach the cecum.

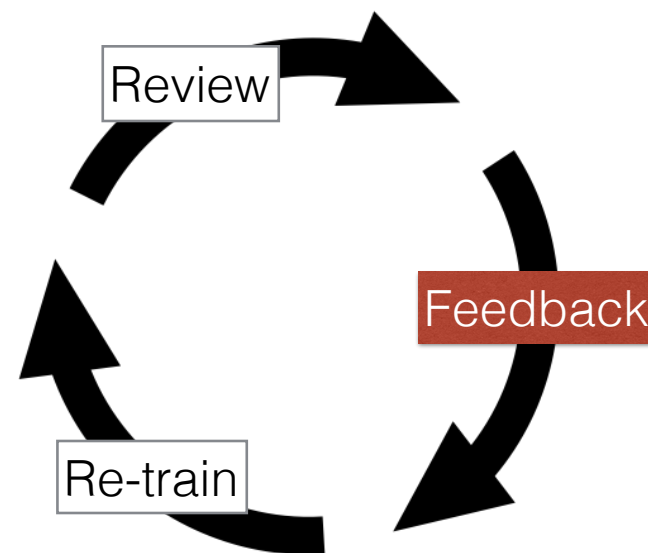
Transverse, descending, sigmoid colon, and rectum appeared normal. Troflexed views in the rectum revealed no abnormalities. The instrument was completely withdrawn from the patient and the procedure was completed.



True indicators  
tubular (0.14)  
adenoma (0.12)  
polypectomy (0.08)  
part (0.08)  
1a (0.07)

False indicators  
scope  
normal  
anal

# Feedback Loop



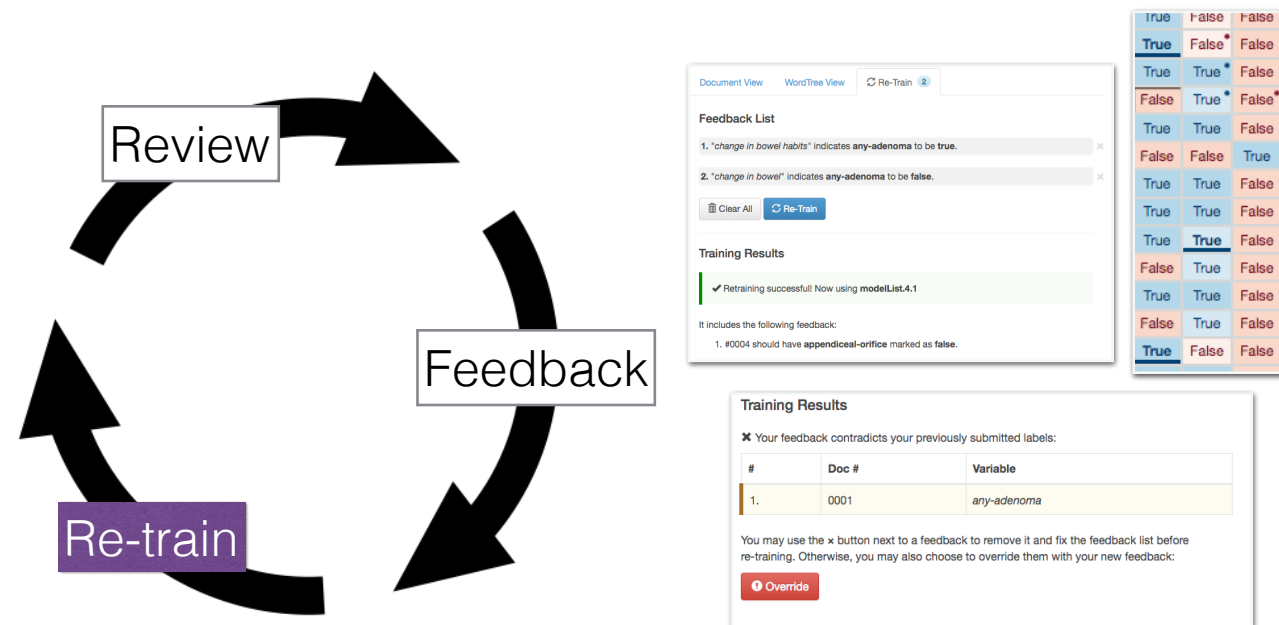
The screenshot displays the Phenol Explorer interface. At the top, a search bar shows 'any-adenoma' with a dropdown arrow and a label 'is marked as false for this record. Label it: [True](#) | [False](#)'. Below this, a table lists various phenotypes with columns: 'any-adenoma', 'any-dissect-site', 'any-peri-dissect-site', 'any-prostatectomy', and 'any-lymphatic-tissue'. The table contains 10 rows of data, with the first row being 'any-adenoma' and the rest being 'any-dissect-site', 'any-peri-dissect-site', 'any-prostatectomy', and 'any-lymphatic-tissue'. The 'any-adenoma' row has a 'True' value in the 'any-adenoma' column and 'True' values in the other columns. The other rows have 'True' values in the 'any-dissect-site' column and 'True' values in the other columns. The 'any-peri-dissect-site' row has a 'True' value in the 'any-peri-dissect-site' column and 'True' values in the other columns. The 'any-prostatectomy' row has a 'True' value in the 'any-prostatectomy' column and 'True' values in the other columns. The 'any-lymphatic-tissue' row has a 'True' value in the 'any-lymphatic-tissue' column and 'True' values in the other columns.

Below the table, a section titled 'change in bowel habits' indicates 'any-adenoma' is 'True | False'. To the right, a 'Doc #' field shows '0312'. Below this, a section titled 'ID#:' shows 'ID#-NUM PROCEDURE: Colonoscopy, diagnostic'. Below this, a section titled 'NAME:' shows 'NAME [AAA, BBB] NAME [CCC: YYY ZZZI, MD]'. Below this, a section titled 'SEX:' shows 'female FEMALE'. Below this, a section titled 'ASSISTANT:' shows 'NAME [XXX, WWW], RN and NAME [VVV, UUU], RN'. Below this, a section titled 'DOB:' shows '05/17/1951' and 'WARD: outpatient'. Below this, a section titled 'AGE:' shows '56' and 'DATE: 03/27/2008'. Below this, a section titled 'ASA CLASS:' shows 'Class II'. Below this, a section titled 'INDICATIONS:' shows '1) change in bowel habits'. Below this, a section titled 'MEDICATION:' shows '200 micrograms Fentanyl IV, 6 milligrams Vers Benadryl IV'.

Below the record details, a section titled 'prep was' indicates 'any-adenoma' is 'True | False'. To the right, a 'Doc #' field shows '0312.0433.0275.0709.0X'. Below this, a section titled 'Document View' shows a word cloud visualization. The word cloud contains the words 'the', 'prep was', 'good', 'excellent', 'adequate', 'moderate', '20 observed', 'for', 'patient', 'poor', 'adequate to the extent', 'somewhat limited in the colon', 'suboptimal', and 'extremely limited'. The word 'prep was' is the largest and most prominent, followed by 'good', 'excellent', 'adequate', 'moderate', '20 observed', 'for', 'patient', 'poor', 'adequate to the extent', 'somewhat limited in the colon', 'suboptimal', and 'extremely limited'.



# Feedback Loop



# User Studies

## **Goals**

- Get a global view of the usability factors
- Gain insight into the design problems leading to any confusion or error

# User Studies

## Goals

- Get a global view of the usability factors
- Gain insight into the design problems leading to any confusion or error

## Participants

- Familiar with clinical texts - *5 Clinicians and clinical researchers*
- Knowledge of colonoscopies and related procedures - *Performed colonoscopies; had a MD degree or more*

# User Studies

## **Subjective Comments**

- The participants were asked to:
  - review documents using the tool and,
  - revise NLP models by providing feedback
- We followed the “*Think aloud*” protocol to record their comments and reactions to the tool
- Survey questionnaires before and after the study

# User Studies

## **System Usability Scale (SUS)**

- We used the System Usability Scale:
  - 10-questions on a 5-point Likert scale
- The average SUS score was 70.5 out of 100.
- Lower scores for questions about the learnability of the tool as we were able to provide a very short walkthrough of the interface to our participants due to time constraints

Brooke (1996), *System Usability Scale (SUS)*

45

Who knew physicians had so little time on their hands ;) The participants acknowledged we could have done better if we were to spend more time on it.

# Overview

- Problem
- Research Goals
- Proposed Solution
- Methods and Evaluation
- **Next Steps and Challenges**

# Next Steps

- We have built a prototype tool for interactive machine learning on clinical text
- Can support users without prior NLP experience to build models interactively
- Next steps involve incorporating user recommendations and conducting an empirical evaluation

# Next Steps

Hypotheses for empirical evaluation:

**H1:** An interactive tool will facilitate the review of clinical text and building NLP models

**H2:** Visual presentation and feedback user interface-components will allow quicker and/or more accurate completion of the task of building NLP models

**H3:** Manual review using the tool may enable rapid convergence on highly accurate models even when starting out with smaller training sets



# Challenges

- Determining the appropriate study design for the empirical study
- Recruiting 20-30 physicians may not be feasible
- Also, harder to differentiate between control conditions

but otherwise the results may not statistically compelling

# Other Limitations

- We have considered only binary variables in this prototype tool
- How do we support more complex NLP applications such as finding *drug-drug interactions*?

# Acknowledgements

- I thank my advisor Dr. Harry Hochheiser, and collaborators on the project, Phuong Pham, Dr. Wendy Chapman, Dr. Rebecca Hwa and Dr. Janyce Wiebe
- Also grateful to our user study participants, and Dr. Ateev Mehrotra for providing the colonoscopy dataset
- This research is supported by NIH grant 5R01LM010964
- Demo video: <http://vimeo.com/trivedigaurav/emr-demo>