MOVING BEYOND SUPERVISED REALM

Regina Barzilay
CSAIL, MIT

NLP@ICLR

DEEP BIAFFINE ATTENTION FOR NEURAL DEPENDENCY PARSING

REASONET: LEARNING TO STOP READING IN MACHINE COMPREHENSION

VOCABULARY SELECTION STRATEGIES FOR NEURAL MACHINE TRANSLATION

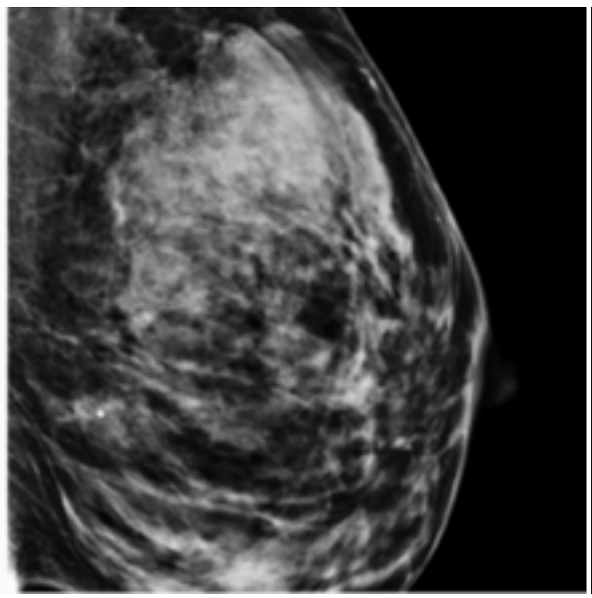
DEEP CHARACTER-LEVEL NEURAL MACHINE TRANSLATION BY LEARNING MORPHOLOGY

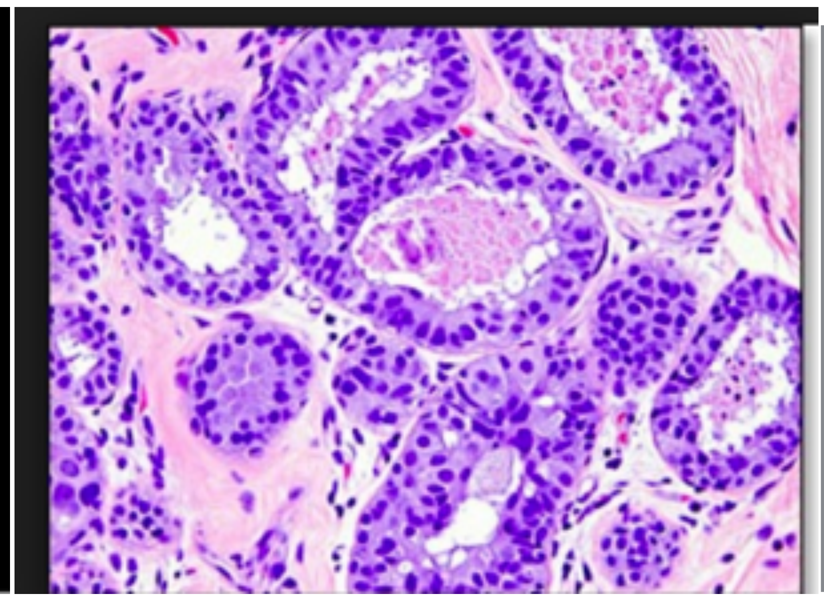
ITERATIVE REFINEMENT FOR MACHINE TRANSLATION

A CONVOLUTIONAL ENCODER MODEL FOR NEURAL MACHINE TRANSLATION

REFERENCE-AWARE LANGUAGE MODELS

Modeling Disease Progression





AST	8~ 40 U/L
ALT	5~ 40 U/L
LDH	50~400 U/L
ALP	80~280 U/L
γ-GTP	0~ 50 U/L
ZTT	4~ 12 KU
TTT	0~ 5 KU
Total bilirubin	0.4~0.9 mg/dℓ
Direct bilirubin	0~0.4 mg/dℓ
Indirect bilirubin	0.4~0.5 mg/dℓ
Total protein	6.0~9.0 g/dℓ

Clinical Prediction

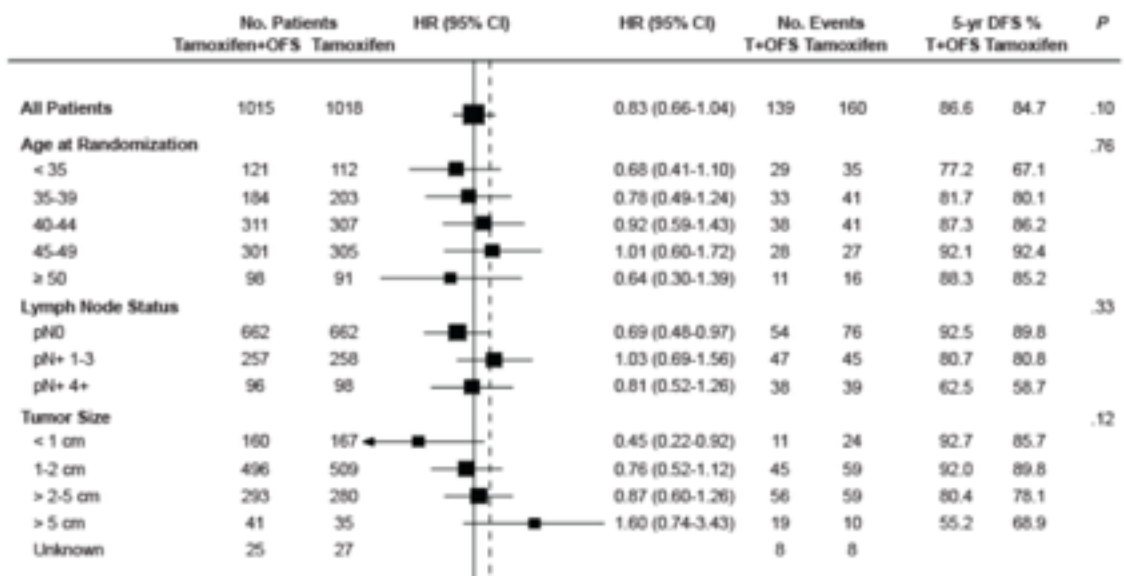
Predict Recurrences, Sensitivity to Treatment, Population at Risk

Game of 20 Questions

What are the chances of recurrence?

Treatment A vs Treatment B?

Drug A vs Drug B?



Answers:

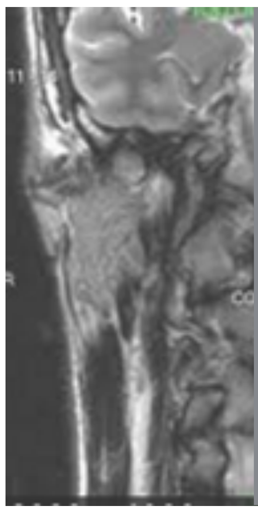
- Check this clinical study, but it is [old, inconclusive,…]
- Your decision at the end

Today, almost all of our cancer treatment insights come from a tiny subset of clinical trial patients. In the United States, 1.7 million people are diagnosed each year with cancer, but only 3% enroll in clinical trials. To improve care for every patient, we need insights from the other 97% of people receiving cancer care.

ASCO
CANCER•LINQ™
Learning Intelligence Network for Quality

Data Science Perspective on Clinical Research

- Abstract clinical records into a database



Accession Number: Report Status: Updated
Type: Surgical Pathology
Pathology Report:
CASE: PATIENT:
Date Taken : 2/12/2008 Source Care Unit: Same Day Surgery Unit

Path Subspecialty Service:Breast-1 Results To: Signed Out by:

Results
CLINICAL DATA right breast CA

FINAL DIAGNOSIS:
BREAST (RIGHT), EXCISION:
1. INVASIVE DUCTAL CARCINOMA WITH CALCIFICATIONS, SEE TABLE #1
2. DUCTAL CARCINOMA *IN-SITU* WITH CALCIFICATIONS
3. LOBULAR NEOPLASIA (ATYPICAL LOBULAR HYPERPLASIA)
4. HEALING BIOPSY SITE

TABLE OF PATHOLOGICAL FINDINGS #1
INVASIVE CARCINOMA
TUMOR SIZE: 1.2 x 1.1 x 0.9 cm (gross measurement)
GRADE: 2
LYMPHATIC VESSEL INVASION: Not identified
BLOOD VESSEL INVASION: Notidentified
MARGIN OF INVASIVE CARCINOMA: Focally positive at the anterior and medial margins of the specimen
LOCATION OF DUCTAL CARCINOMA *IN-SITU*: Within and beyond the region of the mass
GRADE OF DUCTAL CARCINOMA *IN-SITU*:2
MARGIN OF DUCTAL CARCINOMA *IN-SITU*: The distances to all margins measure 0.2 cm or greater
STAINS FOR RECEPTORS: Requested on block B3

ID	AGE	RACE	STUDY	PROC	BIRTHS	MA_AGE	ASSESS	DENSITY	FINDING	FINDING T
9527	78	2	6/12/06	BIDXU-L	0		P	3	CALCS	N
32875	56	1	7/11/06	BIDXB-B	0		N	3		
2247	72	1	4/12/06	BIDXU-R	0		N	3		
45521	61	1	3/30/06	BIDXB-B	0		B	3	CALCS	S
48987	41	1	4/5/06	BIDXB-B	0		P	3	CALCS	N
4179	67	1	5/12/06	BIDXB-B	0		P	2	CALCS	N
26300	59	1	3/31/06	BIDXU-L	0		N	3		
67960	64	1	4/7/06	BIDXU-R	0		P	3	MASS	O
43283	61	W	7/21/06	BIDXB-B	0		B	3		
43319	51	1	4/7/06	BIDXB-B	0		N	3		

- Run multivariate analysis to identify the correlation between features and outcomes

Standard data analysis methodology in modeling disease progression, treatment efficacy, and diagnosis

Manually Constructed Databases Rule!

- Seriously labour intensive

- Fraught with unknown values (see # of births) and errors

ID	AGE	RACE	STUDY	PROC	BIRTHS	MA_AGE	ASSESS	DENSITY	FINDING	FINDING T
9527	78	2	6/12/06	BIDXU-L	0		P	3	CALCS	N
32875	56	1	7/11/06	BIDXB-B	0		N	3		
2247	72	1	4/12/06	BIDXU-R	0		N	3		
45521	61	1	3/30/06	BIDXB-B	0		B	3	CALCS	S
48987	41	1	4/5/06	BIDXB-B	0		P	3	CALCS	N
4179	67	1	5/12/06	BIDXB-B	0		P	2	CALCS	N
26300	59	1	3/31/06	BIDXU-L	0		N	3		
67960	64	1	4/7/06	BIDXU-R	0		P	3	MASS	O
43283	61	W	7/21/06	BIDXB-B	0		B	3		
43319	51	1	4/7/06	BIDXB-B	0		N	3		

Can we automate database construction from raw text?

Looking for Ways to Contribute

Predicting Progression of Atypical Lesions

Clinical goal: use chemo prevention to modify the risk of atypical breast lesions

Method: observe lesion progression over time

- some lesion types are rare
- analysis requires large corpus of pathology reports

Kevin S. Hughes, M.D., FACS
Surgical Director, Breast Screening
Co-Director, Avon Comprehensive Breast Evaluation Center
Massachusetts General Hospital

Parsing Pathology Reports into Database

Pathology Report: REMOVED_ACCESSION_ID
ACCESSIONED ON: REMOVED_DATE
CLINICAL DATA: Carcinoma **right breast**.
*** FINAL DIAGNOSIS ***
LYMPH NODE (SENTINEL), EXCISION
(REMOVED_CASE_ID): METASTATIC
CARCINOMA IN 1 OF 1 LYMPH NODE.
NOTE: The metastatic deposit spans 0.19cm and is identified on H&E and cytokeratin immunostains. A second cytokeratin-positive but cauterized focus likely also represents metastatic tumor (<0.1cm). There is **no evidence of extranodal extension**.
BREAST (RIGHT), EXCISIONAL BIOPSY
(REMOVED_ACCESSION_ID :
REMOVED_CASE_ID -B): **INVASIVE DUCTAL CARCINOMA** (SEE TABLE #1). **DUCTAL CARCINOMA IN-SITU**, GRADE 1. **ATYPICAL DUCTAL HYPERPLASIA**. **LOBULAR NEOPLASIA (ATYPICAL LOBULAR HYPERPLASIA)**.
TABLE OF PATHOLOGICAL FINDINGS #1

Name	**Extraction**
Breast Side	Right
Ductal Carcinoma in Situ	Present
Invasive Lobular Carcinoma	Absent
Invasive Ductal Carcinoma	Present
Cancer	Present
Lobular Carcinoma in Situ	Absent
Atypical Ductal Hyperplasia	Present
Atypical Lobular Hyperplasia	Present
Lobular Neoplasia	Present
Flat Epithelial Atypia	Absent
Blunt Adenosis	Absent
Atypia	Present
Positive Lymph Nodes	Present
Extracapsular Axillary Nodal Extension	Absent
Isolated Cancer Cells in Lymph Nodes	Absent
Lymphovascular Invasion	Absent
Blood Vessel Invasion	Absent
Estrogen Receptor Status	Positive
Progesterone Receptor Status	Positive
HER 2 (FISH) Status	Unknown

Categorization of Reports

22 ways to say LCIS

Lobular Carcinoma In-Situ
Lobular Carcinoma In Situ
Lobular Carcinoma-In-Situ
Lobular Carcinoma In -Situ
Lobular Cardinoma In-Situ
In-Situ Carcinoma Is Of Lobular
In-Situ Carcinoma To Be Lobular
In-Situ Carcinoma With Lobular
In-Situ Carcinoma With Both Lobular
In Situ Carcinoma And Atypical Lobular
In Situ Carcinoma Has Ductal And Lobular
In-Situ Carcinoma With Ductal And Lobular

In Situ Carcinoma Is Interpreted As Lobular
In Situ Carcinoma With Ductal And Lobular
In Situ Carcinoma May Be Ductal Or Lobular
In Situ Carcinoma With Features Of Both Lobular
In-Situ Carcinoma Has Some Features Of Lobular
In-Situ Carcinoma With Both Ductal And Lobular
In-Situ Carcinoma With Features Of Both Lobular
In-Situ Carcinoma With Mixed Ductal And Lobular
In-Situ Carcinoma Displays Both Ductal And Lobular
In-Situ Carcinoma Showing Both Ductal And Lobular
In-Situ Carcinoma Are Small And Distinction Of Lobular
In-Situ Carcinoma With Features Of Ductal And Lobular
In Situ Carcinoma Demonstrates Both Ductal And Lobular

slide from K. Hughes

NLP Leads to Important Clinical Finding!

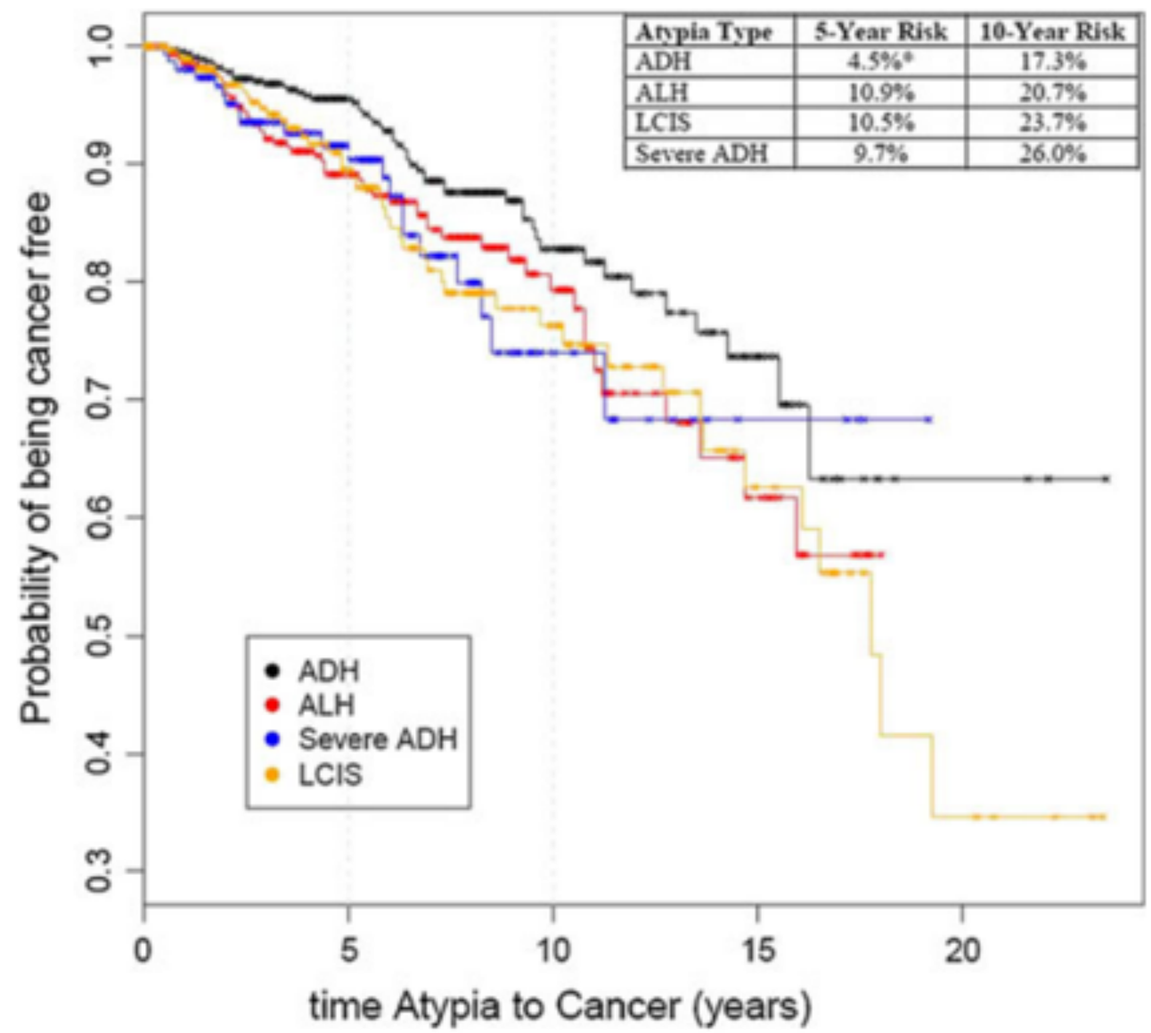


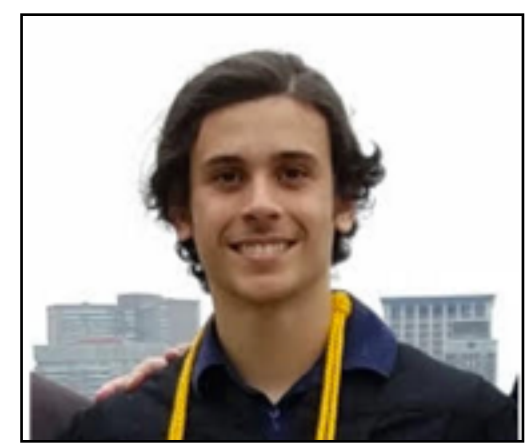
Fig. 1 Estimated 5- and 10-year breast cancer risks based on atypia type for the no chemoprevention group. *Significantly fewer predicted breast cancers at 5 years with ADH ($p = 0.036$)

The role of chemoprevention in modifying the risk of breast cancer in women with atypical breast lesions

NLP can Help … But …

NOW NORMAL WAY

- Method: train per category classifier based on ngram features
- Training size: 6,000 reports
- Predicts 20 categories, ranging from atypias to tumor markers
- Accuracy: 98.6% (evaluated by two MDs on 500 reports)
- Size: 92K reports, 50K patients, over 30 years
- Development Time: two weeks



Adam Yala, MIT

Rule-Based IE?!!

- Common in Medical and Bio Informatics

Using Natural Language Processing to Improve Efficiency of Manual Chart Abstraction in Research: The Case of Breast Cancer Recurrence

David S. Carrell,* Scott Halgrim, Diem-Thy Tran, Diana S. M. Buist, Jessica Chubak, Wendy W. Chapman, and

Marco Antonio Valenzuela-Escarcega, Gustave Hahn-Powell, Mihai Surdeanu, Thomas Hicks: **A Domain-independent Rule-based Framework for Event Extraction.** ACL (System Demonstrations) 2015: 127-132

- And Beyond …

Rule-Based Information Extraction is Dead! Long Live Rule-Based Information Extraction Systems!

Laura Chiticariu, Yunyao Li, Frederick Reiss
EMNLP, *pp. 827-832*, 2013

Aspect Transfer

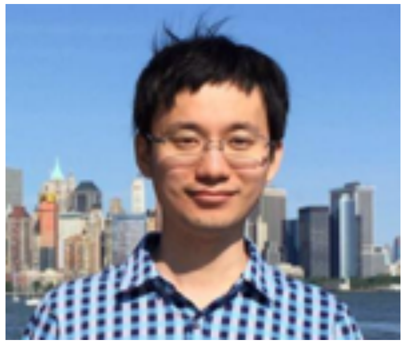
Pathology report:

FINAL DIAGNOSIS: BREAST (LEFT) … INVASIVE CARCINOMA Tumor size: num x num x num cm Grade: 3. Lymphatic vessel invasion: Not identified. Blood vessel invasion: Suspicious. Margin of invasive carcinoma …

Diagnosis results:

IDC: Positive LVI: Negative

Transfer: Source: IDC → Target: LVI



Yuan Zhang

Tommi Jaakkola

Multi-Aspect Transfer

***Same** report; **Different** key sentences*

Source Aspect: IDC

Target Aspect: LVI

FINAL DIAGNOSIS: BREAST (LEFT) … INVASIVE CARCINOMA Tumor size: num x num x num cm Grade: 3. Lymphatic vessel invasion: Not identified. Blood vessel invasion: Suspicious. Margin of invasive carcinoma …

What to Transfer?

FINAL DIAGNOSIS: BREAST (LEFT) … INVASIVE CARCINOMA Tumor size: num x num x num cm Grade: 3. Lymphatic vessel invasion: Not identified. Blood vessel invasion: Suspicious. Margin of invasive carcinoma …

FINAL DIAGNOSIS: BREAST (RIGHT) The specimen examined by X. DCIS: Not Identified. ….

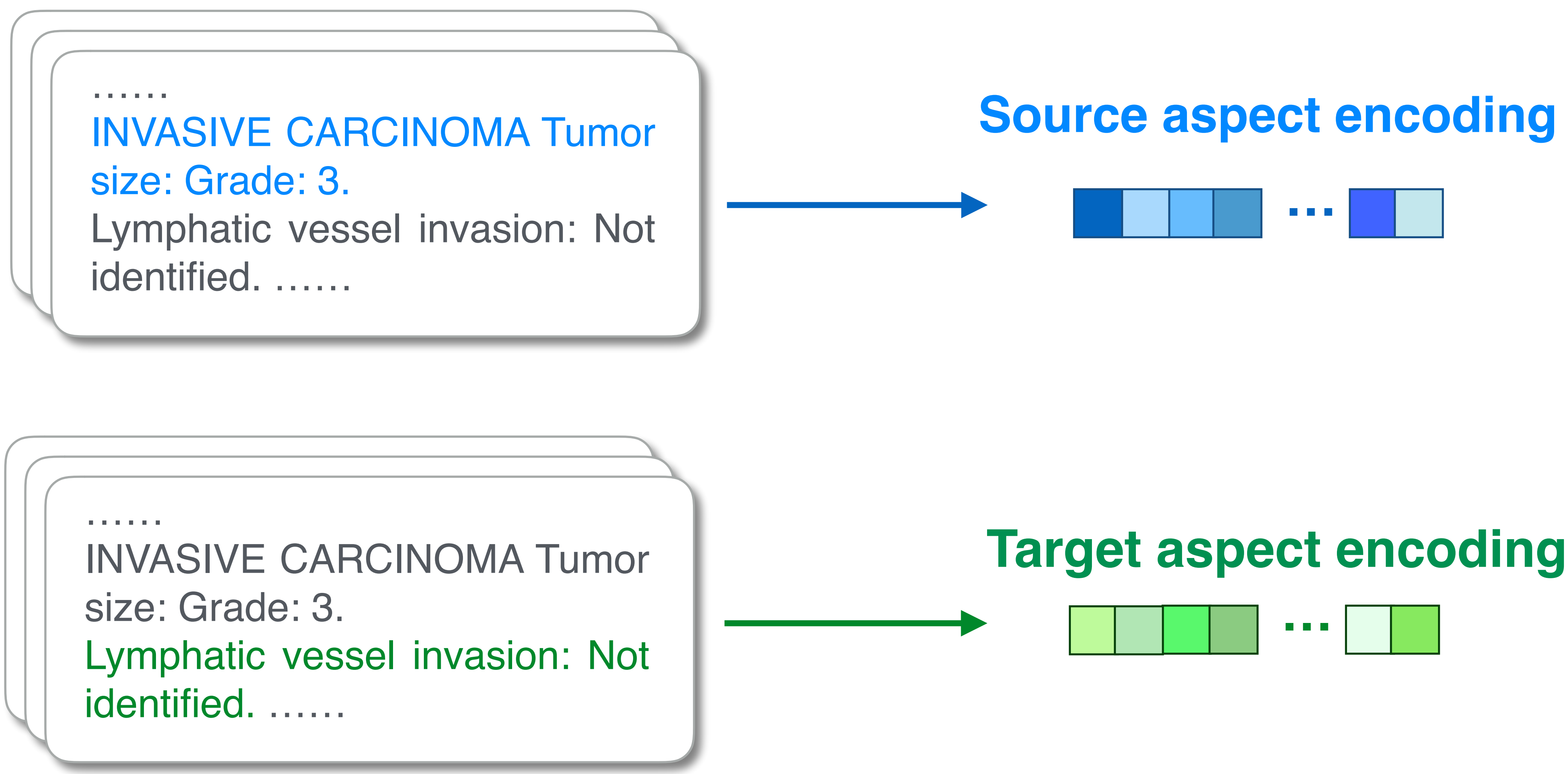
Available Supervision

	Source	Target
Labeled Data	✓	✗
Unlabeled Data	✓	✓
Relevance Rules	✓	✓

- Relevance rules: common names of aspects
  - ALH: Atypical Lobular Hyperplasia, ALH
  - IDC: Invasive Ductal Carcinoma, IDC

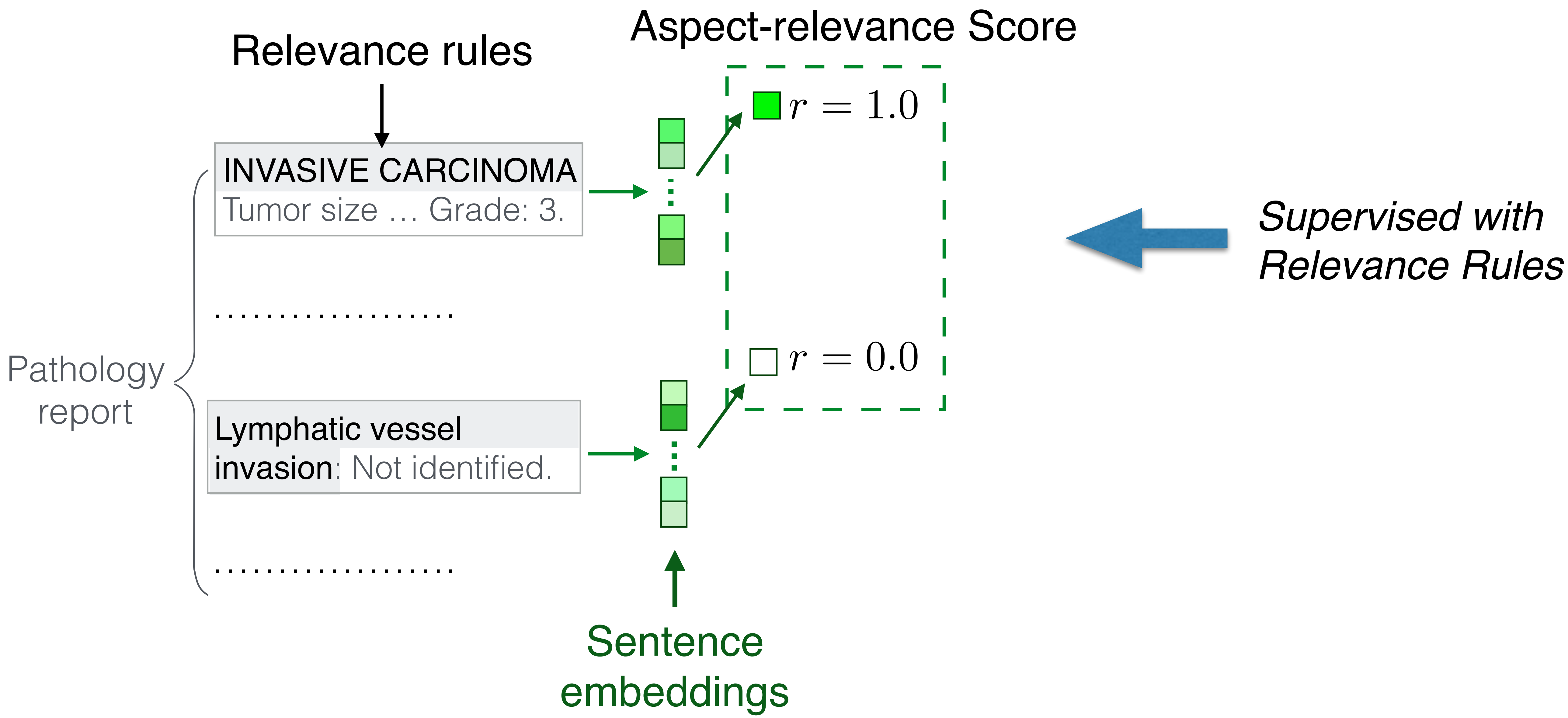
Key Idea: Aspect-driven Encoding

- Learn differential representations for different tasks from the same input
- Leverage relevance information to learn to identify key fragments



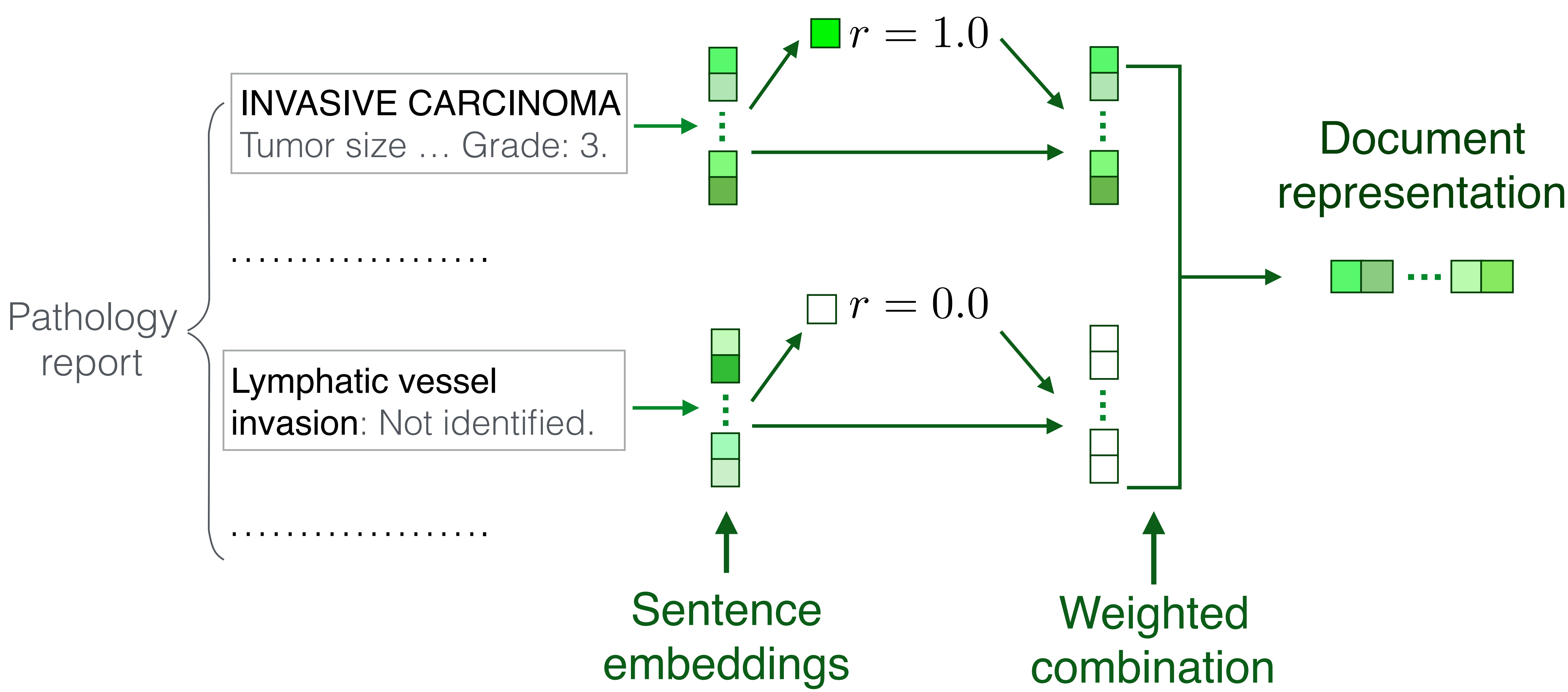
Aspect-relevance Prediction

- Make prediction based on sentence embeddings
- Train on relevance rules (e.g., names of IDC, LVI)



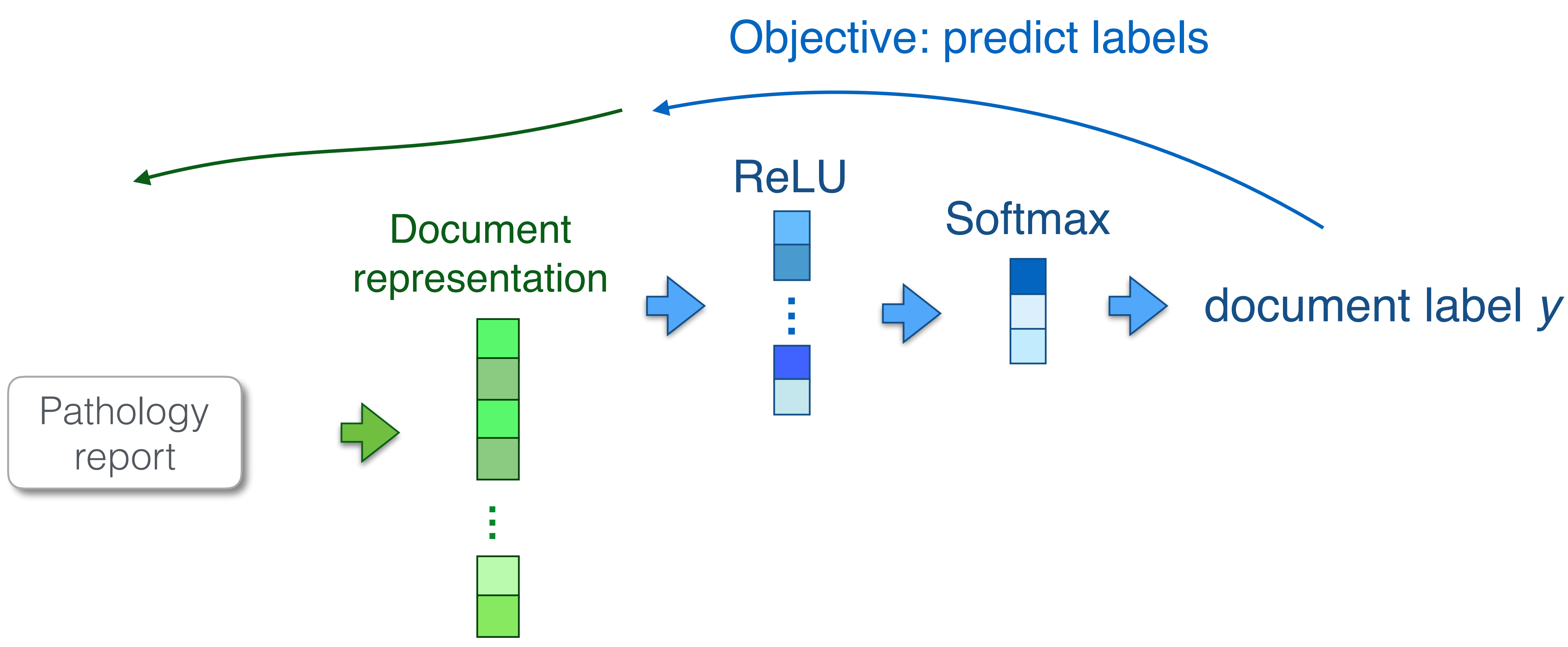
Document Encoding

- Combine sentence vectors based on relevance weights



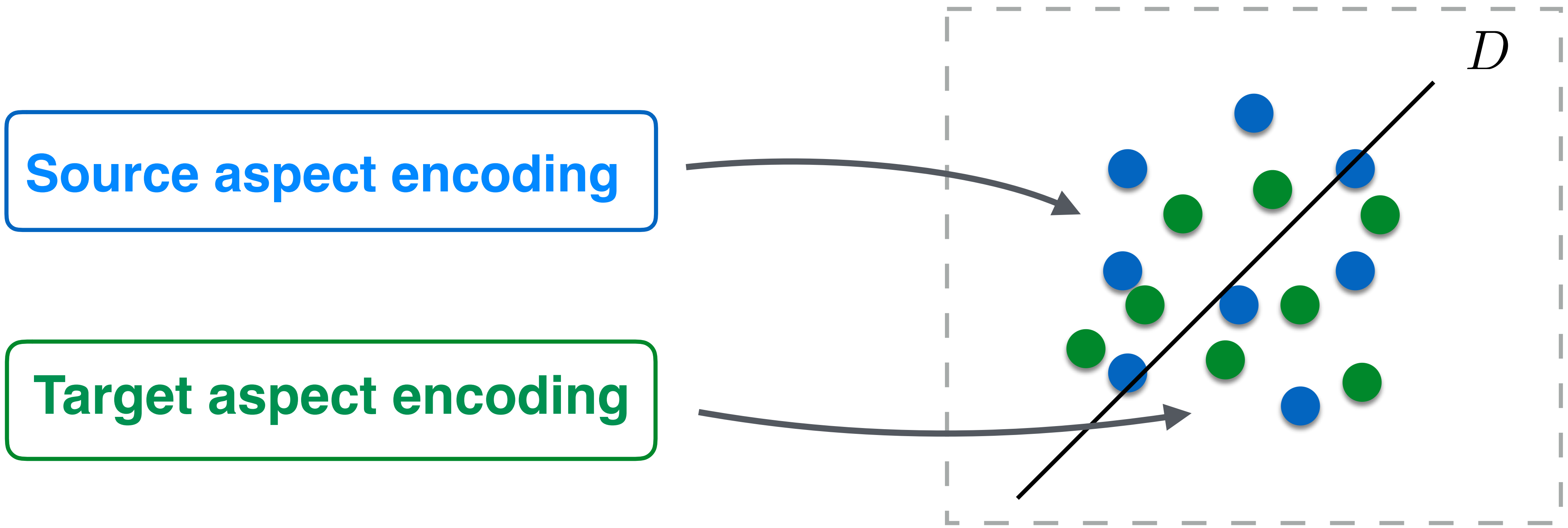
Primary Label Predictor

- Shared for both source and target aspects
- Train on labeled data in the source aspect



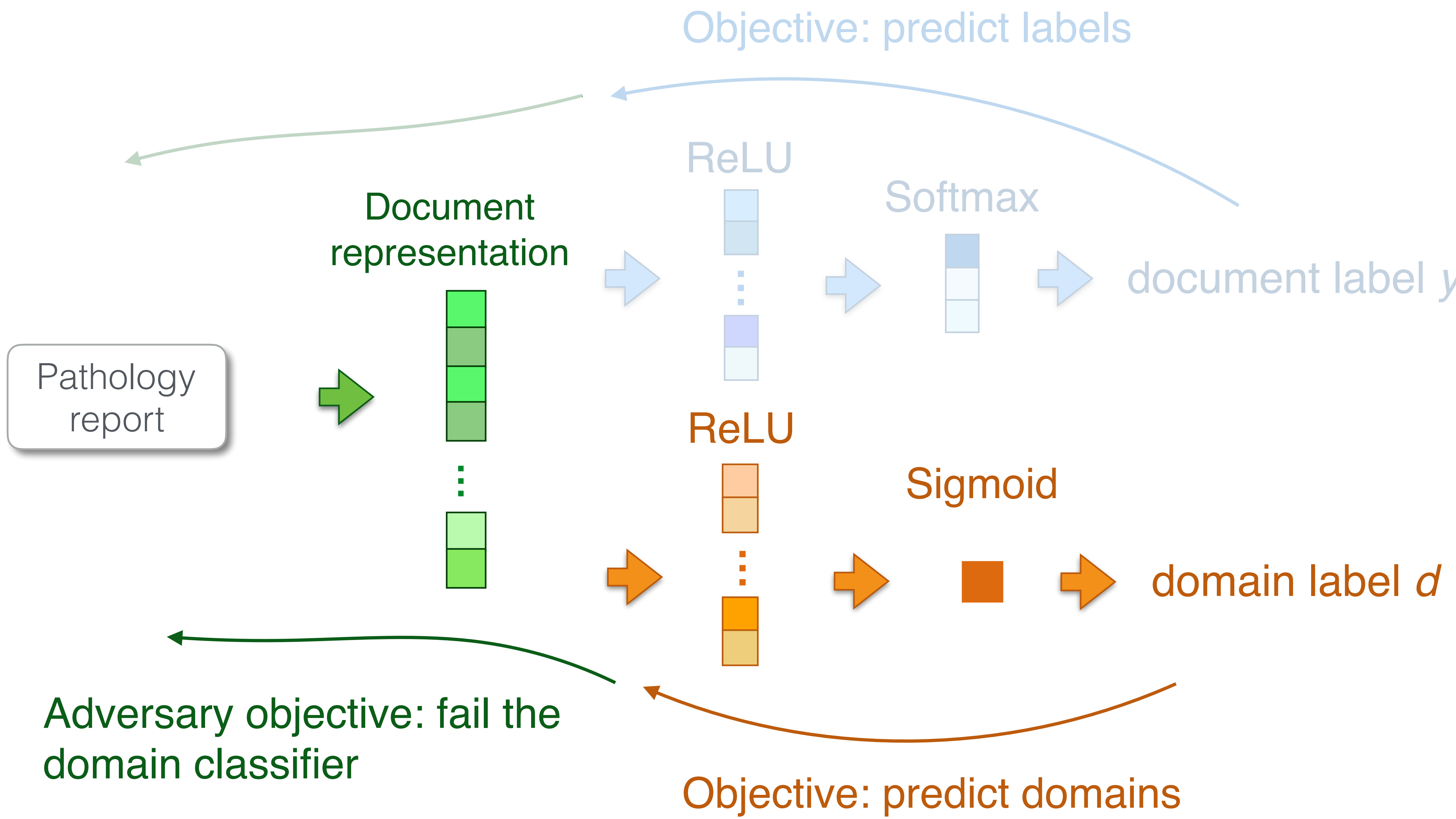
Key Idea: Domain-Adversarial Training

- Use domain-adversarial for learning invariant representations
  - Objective: Not separable by the domain

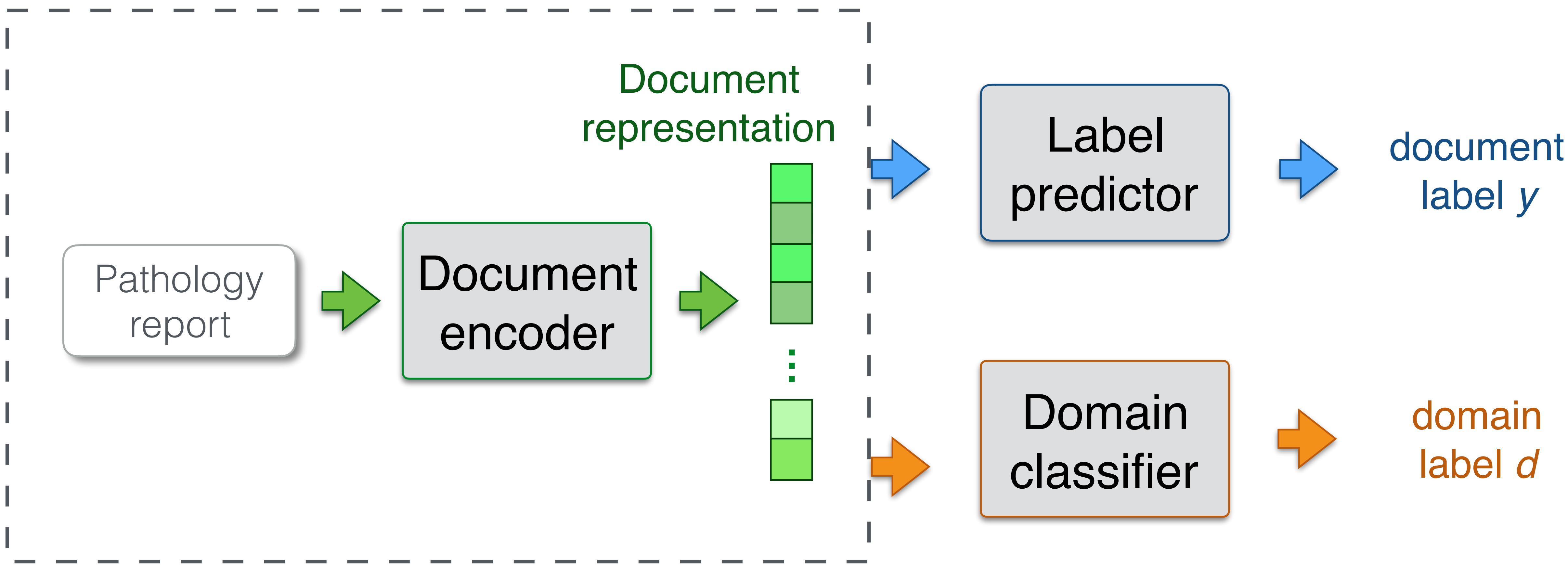


Domain Classifier and Adversary

- Encourage learning domain-invariant features
- Train on both labeled and unlabeled data

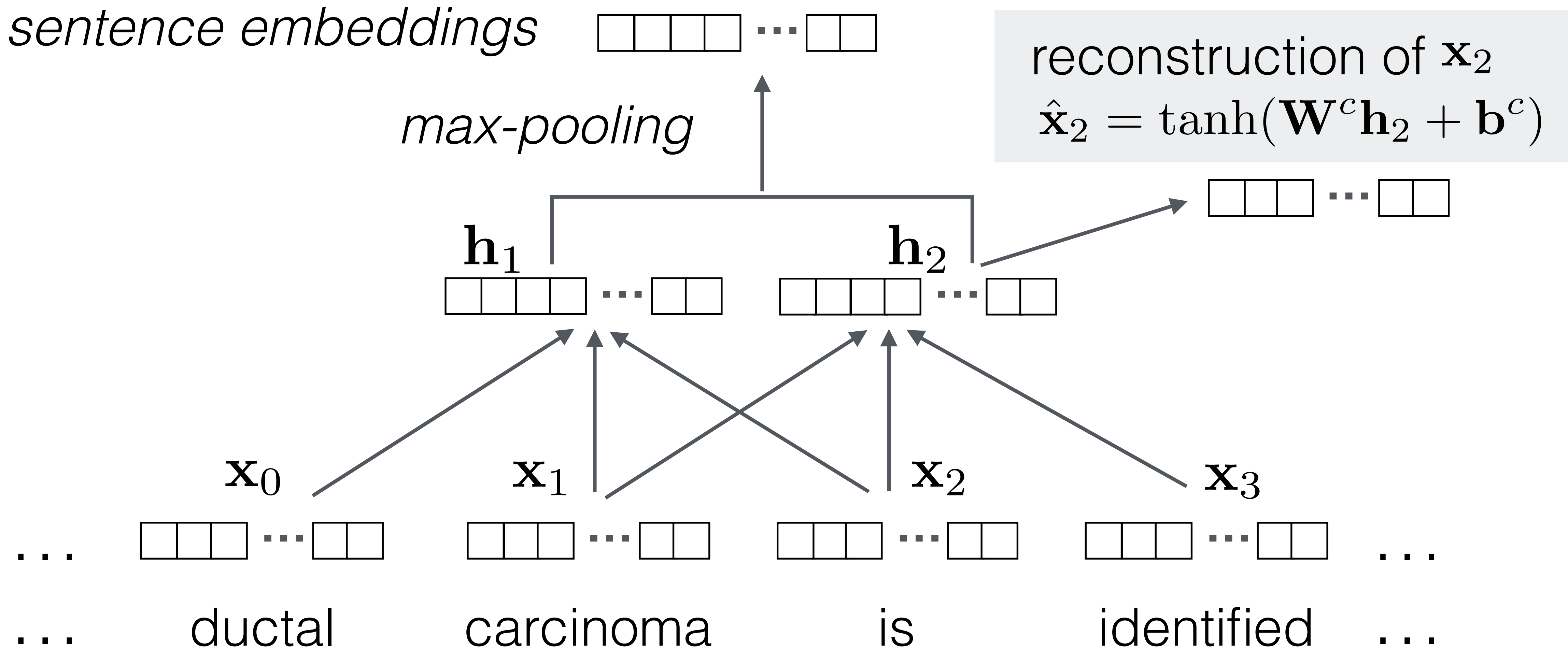


Model Structure



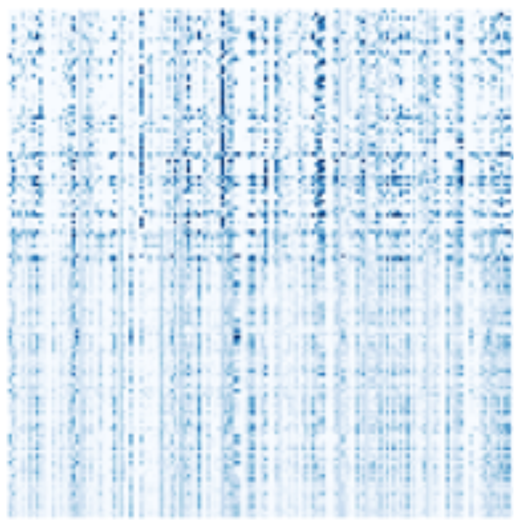
Sentence Embedding

- Apply a CNN to each sentence
- Keep rich word-level information by reconstruction

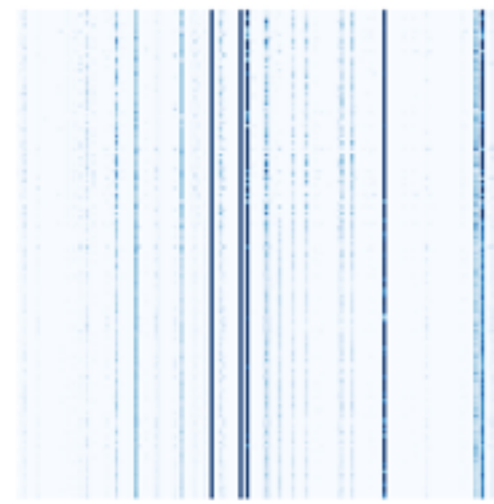


Impact of Reconstruction

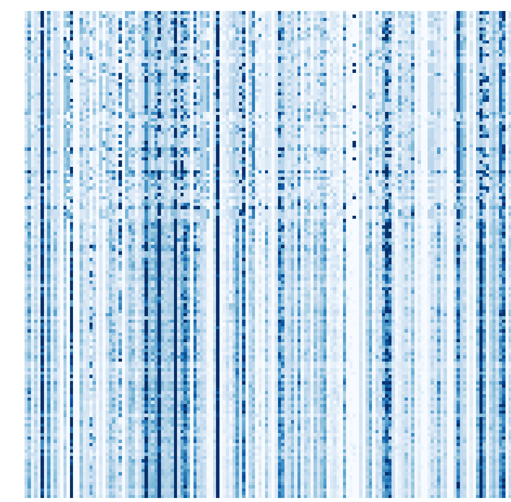
- Heat-map: each row correspond to a vector representation
  - Top: source domain; Bottom: target domain



−adversarial, −reconstruction

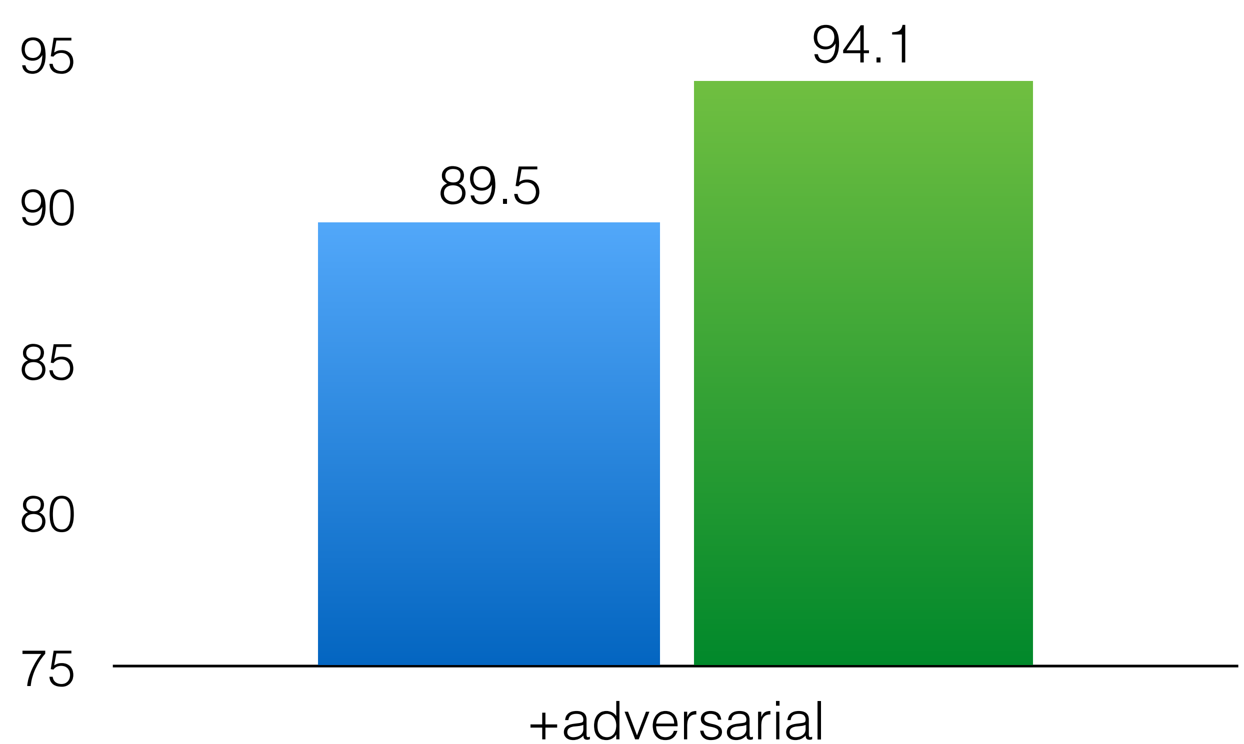


+adversarial, −reconstruction



+adversarial, +reconstruction

- Average accuracy on the pathology dataset



Pathology Dataset

- Aspect-transfer on breast pathology reports

Source: IDC → Target: LCIS

> FINAL DIAGNOSIS: BREAST (LEFT) … INVASIVE DUCTAL CARCINOMA Grade: 3. Lobular Carcinoma In-situ: Not identified. Blood vessel invasion: Suspicious. …

- Statistics and relevance rules:

Aspects	#Labeled	#Unlabeled	Relevance Rules
DCIS	23.8k	96.6k	DCIS, Ductal Carcinoma In-Situ
LCIS	10.7k		LCIS, Lobular Carcinoma In-Situ
IDC	22.9k		IDC, Invasive Ductal Carcinoma
ALH	9.2k		ALH, Atypical Lobular Hyperplasia

✦ 500 reports for testing

Review Dataset

- Domain-transfer on reviews

Source: Restaurant → Target: Hotel

The fries were undercooked and thrown haphazardly into the sauce holder …

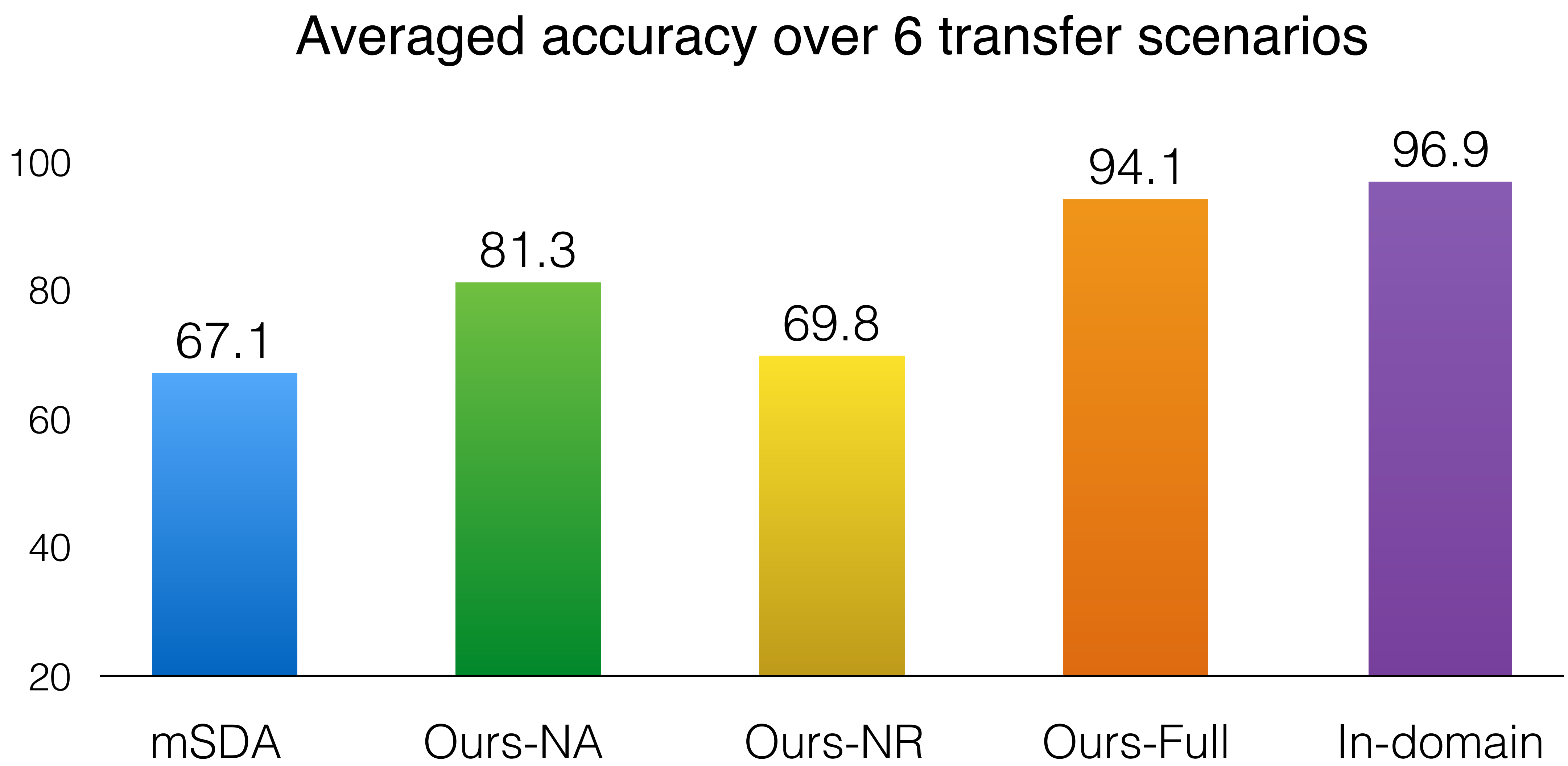
In the second bedroom it literally rained water from above …

- Statistics and relevance rules:

Domains	#Labeled	#Unlabeled	Relevance Rules
Hotel	100k	100k	290 keywords from Wang et al., 2011
Restaurant	-	200k	(only one *overall* aspect)

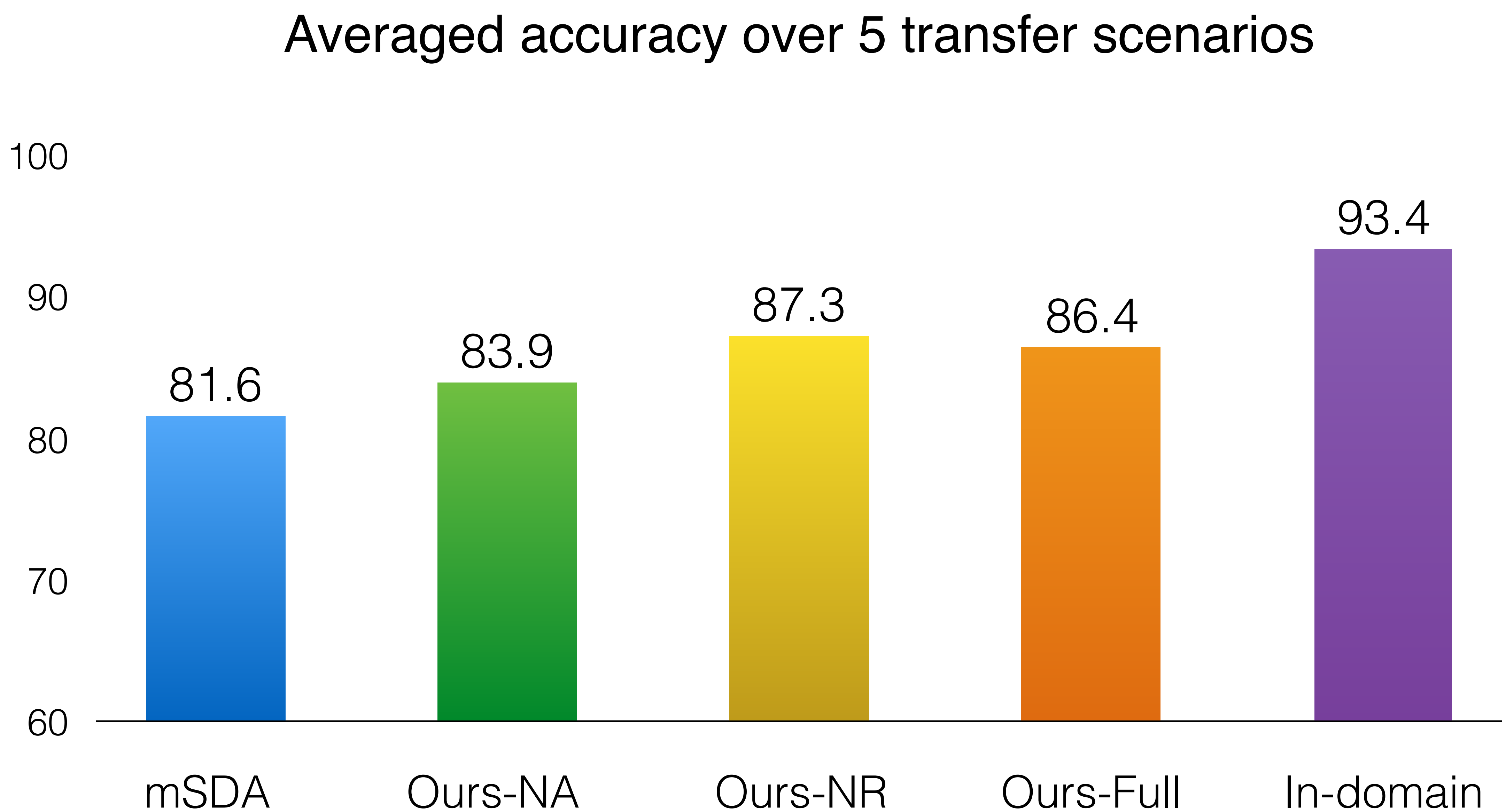
✦ 2k reports for testing

Results on Pathology Dataset



- mSDA: marginalized stacked denoising autoencoder (Chen et al., 2012)
- Ours-NA: our model without adversarial training
- Ours-NR: our model without aspect-relevance scoring
- In-domain: supervised training with in-domain annotations

Results on Review Dataset



Case Study of Learned Representations

Restaurant Reviews

- the fries were **undercooked** and **thrown haphazardly** into the sauce holder . the shrimp was **over cooked** and just **deep fried** . … even the water **tasted weird** .

Nearest Hotel Reviews by **Ours-Full: learns to map domain-specific words**

- the room was **old** . … we did n't like the night shows at all . …
- however , the decor **was just fair** . … in the second bedroom it literally **rained water from above** .

Nearest Hotel Reviews by **Ours-NA: only captures common sentiment phrases**

- rest room in this restaurant is **very dirty** . …
- the only **problem** i had was that … i was very ill with what was suspected to be **food poison**

✦ distance measured by cosine similarity between representations

IE is hard

	Text	# Wounded
Original	**A 2 year old girl and four other people** were wounded in a shooting in West Englewood Thursday night, police said	4

IE is hard - not always!

	Text	# Wounded
Original	**A 2 year old girl and four other people** were wounded in a shooting in West Englewood Thursday night, police said	4
Alternative	The last shooting left **five people wounded**.	5

Easier to extract from alternative sources

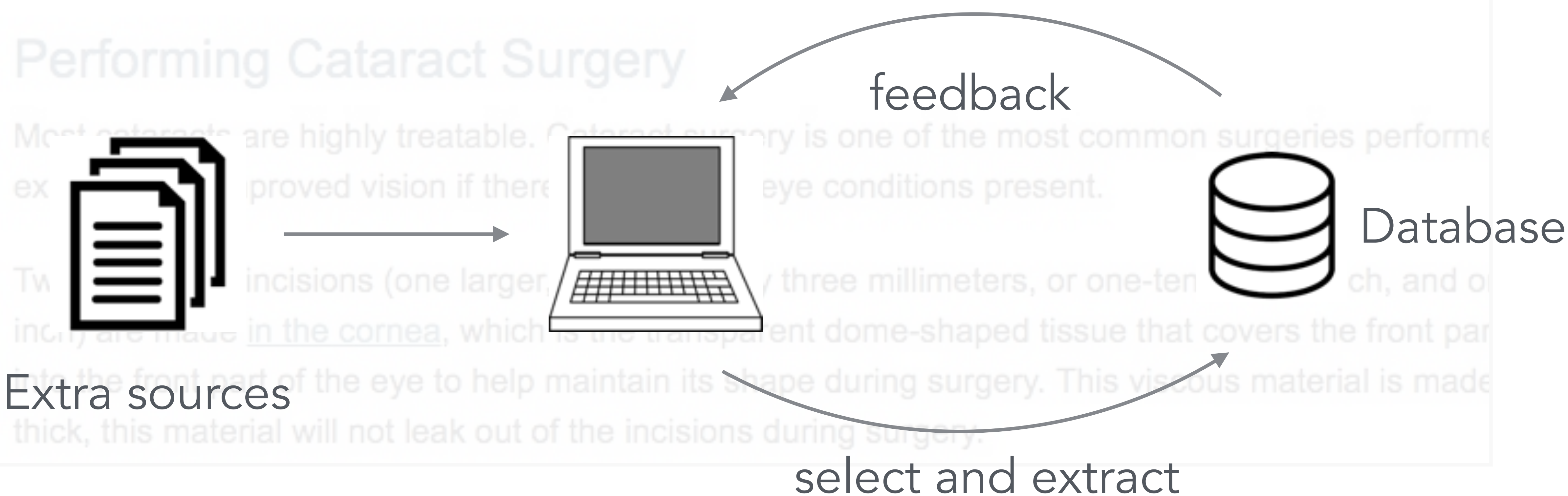
Test in Reading Comprehension

Caffeine significantly reduced ER and cyclin D1 abundance in ER(+) cells. Caffeine also reduced the pAkt levels in both ER(+) and ER(-) cells.

Your information extraction task: Identify cancerogens

Search for Alternative Sources

Automate this process with Reinforcement Learning



Assumptions

1. Access to a basic Information Extraction system
2. Original article to extract target entities from
3. Method to formulate different queries

SECTIONS

DAILY NEWS | NEWS

Crime U.S. World Politics

S.D. dad killed wife, four kids with shotgun before setting house ablaze and killing self: authorities

Scott and Nicole Westerhuis died along with their two sons and two daughters. (NICOLE WESTERHUIS/VIA FACEBOOK)

THE ASSOCIATED PRESS Tuesday, November 3, 2015, 8:59 PM

PLATTE, S.D. — Financial issues appear to have contributed to an educational cooperative business manager's decision to kill his wife and four children with a shotgun before setting the family home ablaze and then shooting himself, South Dakota's attorney general said Tuesday.

Attorney General Marty Jackley released the results of his office's investigation of the September deaths at a news conference in Platte, a few miles north of the burned ruins of the home where the bodies of Scott and Nicole Westerhuis and their children Kailey, Jaeci, Connor and Michael were found.

Acquiring External Evidence

1. Select a query to search for articles on the same event

2. Use base extractor to obtain values for entities of interest


Platte Fire: Westerhuis Family Died in Apparent Murder-Suicide, Officials Say

by HENRY AUSTIN

SHARE

FROM SEPT. 19TH: Westerhuis Neighbor Reacts to Home Fire 0:28

A couple and four children found dead in their burning South Dakota home had been shot in an apparent murder-suicide, officials said Monday.

extract

Shooter: Scott Westerhuis

NumKilled: 6

Location: Platte

3. Reconcile old and new extractions

Shooter: Scott Westerhuis

NumKilled: 4

Location: S.D

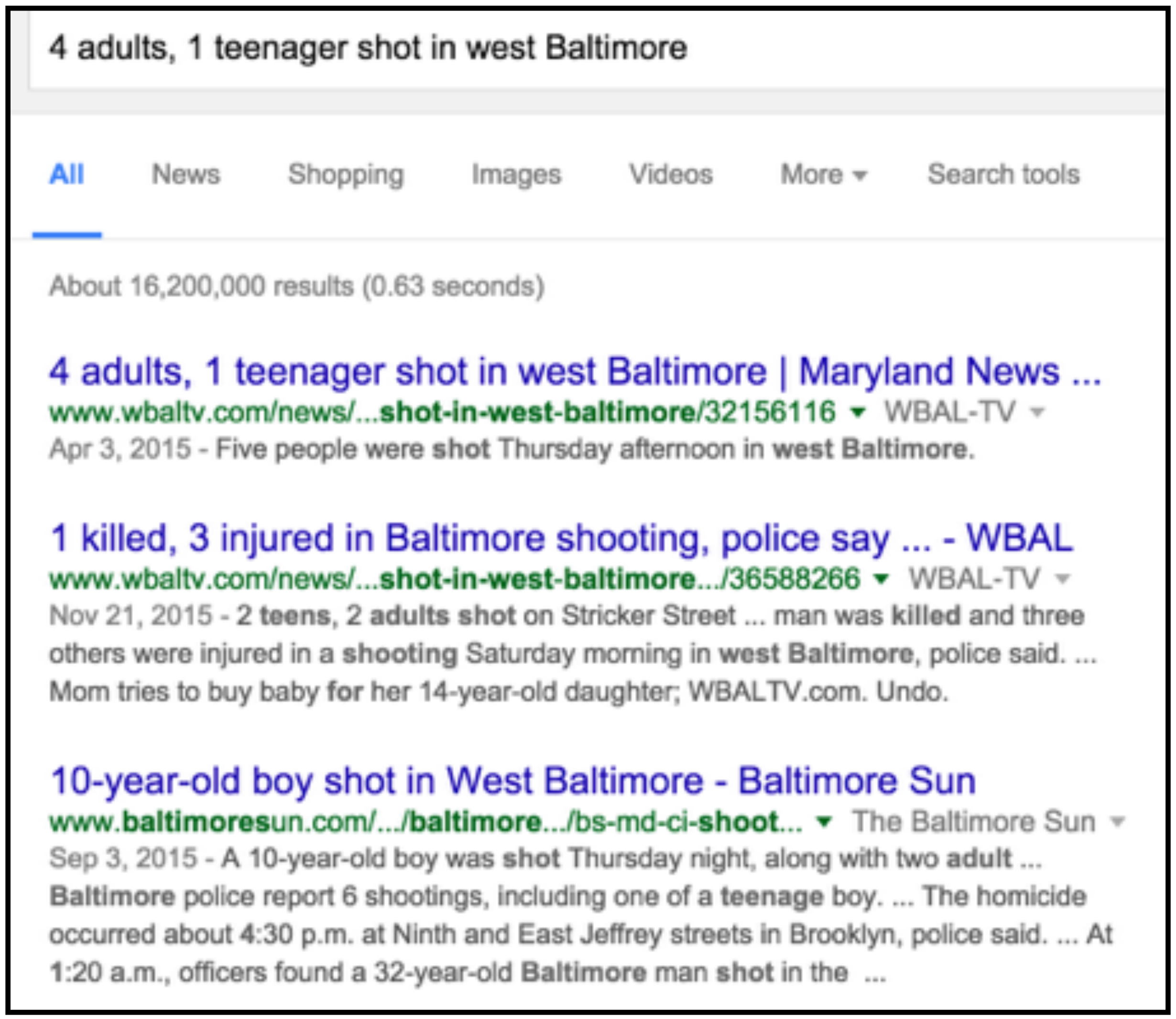
Shooter: Scott Westerhuis

NumKilled: 6

Location: Platte

Challenges

1. Event Coreference


4 adults, 1 teenager shot in west Baltimore

All News Shopping Images Videos More Search tools

About 16,200,000 results (0.63 seconds)

4 adults, 1 teenager shot in west Baltimore | Maryland News ...
www.wbaltv.com/news/...shot-in-west-baltimore/32156116 WBAL-TV
Apr 3, 2015 - Five people were shot Thursday afternoon in west Baltimore.

1 killed, 3 injured in Baltimore shooting, police say ... - WBAL
www.wbaltv.com/news/...shot-in-west-baltimore.../36588266 WBAL-TV
Nov 21, 2015 - 2 teens, 2 adults shot on Stricker Street ... man was killed and three others were injured in a shooting Saturday morning in west Baltimore, police said. ... Mom tries to buy baby for her 14-year-old daughter; WBALTV.com. Undo.

10-year-old boy shot in West Baltimore - Baltimore Sun
www.baltimoresun.com/.../baltimore.../bs-md-ci-shoot... The Baltimore Sun
Sep 3, 2015 - A 10-year-old boy was shot Thursday night, along with two adult ... Baltimore police report 6 shootings, including one of a teenage boy. ... The homicide occurred about 4:30 p.m. at Ninth and East Jeffrey streets in Brooklyn, police said. ... At 1:20 a.m., officers found a 32-year-old Baltimore man shot in the ...

Several irrelevant articles!

2. Reconciling Predictions

Shooter: Scott Westerhuis

NumKilled: 4

Location: S.D

Shooter: Scott Westerhuis

NumKilled: 6

Location: Platte

Varying extractions

Learning through Reinforcement


SECTIONS
DAILY NEWS | NEWS
Crime U.S. World Politics

S.D. dad killed wife, four kids with shotgun before setting house ablaze and killing self: authorities

Scott and Nicole Westerhuis died along with their two sons and two daughters. (NICOLE WESTERHUIS/VIA FACEBOOK)

THE ASSOCIATED PRESS Tuesday, November 3, 2015, 8:59 PM

PLATTE, S.D. — Financial issues appear to have contributed to an educational cooperative business manager's decision to kill his wife and four children with a shotgun before setting the family home ablaze and then shooting himself, South Dakota's attorney general said Tuesday.

Attorney General Marty Jackley released the results of his office's investigation of the September deaths at a news conference in Platte, a few miles north of the burned ruins of the home where the bodies of Scott and Nicole Westerhuis and their children Kailey, Jaeci, Connor and Michael were found.

extract →

Shooter: Scott Westerhuis

NumKilled: 4

Location: S.D

Learning through Reinforcement


DAILY NEWS | NEWS

Crime U.S. World Politics

S.D. dad killed wife, four kids with shotgun before setting house ablaze and killing self: authorities

THE ASSOCIATED PRESS Tuesday, November 3, 2015, 8:59 PM

PLATTE, S.D. — Financial issues appear to have contributed to an educational cooperative business manager's decision to kill his wife and four children with a shotgun before setting the family home ablaze and then shooting himself, South Dakota's attorney general said Tuesday.

Attorney General Marty Jackley released the results of his office's investigation of the September deaths at a news conference in Platte, a few miles north of the burned ruins of the home where the bodies of Scott and Nicole Westerhuis and their children Kailey, Jaeci, Connor and Michael were found.

Scott and Nicole Westerhuis died along with their two sons and two daughters. (NICOLE WESTERHUIS/VIA FACEBOOK)

extract →

Shooter: Scott Westerhuis

NumKilled: 4

Location: S.D

search →


Platte Fire: Westerhuis Family Died in Apparent Murder-Suicide, Officials Say

by HENRY AUSTIN

SHARE

FROM SEPT. 19TH: Westerhuis Neighbor Reacts to Home Fire 0:28

A couple and four children found dead in their burning South Dakota home had been shot in an apparent murder-suicide, officials said Monday.

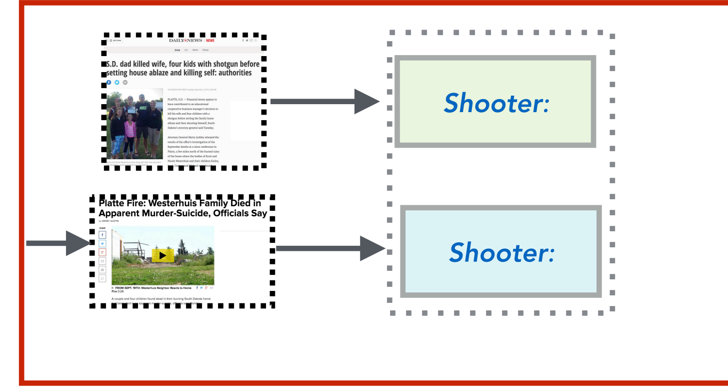
extract →

Shooter: Scott Westerhuis

NumKilled: 6

Location: Platte

State



RL: State

Current Values	Conf
Shooter: Scott Westerhuis	0.3
NumKilled: 4	0.2
Location: S.D.	0.1

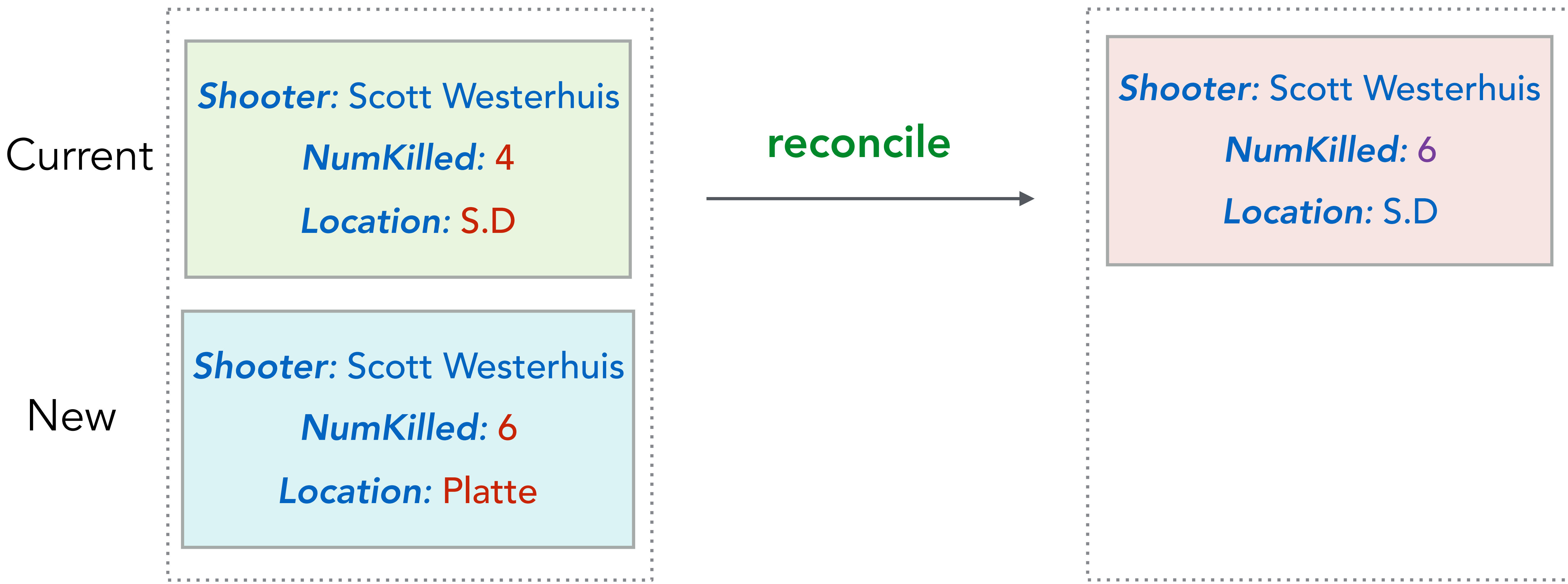
New Values	Conf
Shooter: Scott Westerhuis	0.4
NumKilled: 6	0.6
Location: Platte	0.3

< 0.3, 0.2, 0.1, 0.4, 0.6, 0.3, 1, 0, 0, 0.65, 0, 0, 1, … , 0, 0 >

currentConf newConf matches docSim context

Real-valued vector

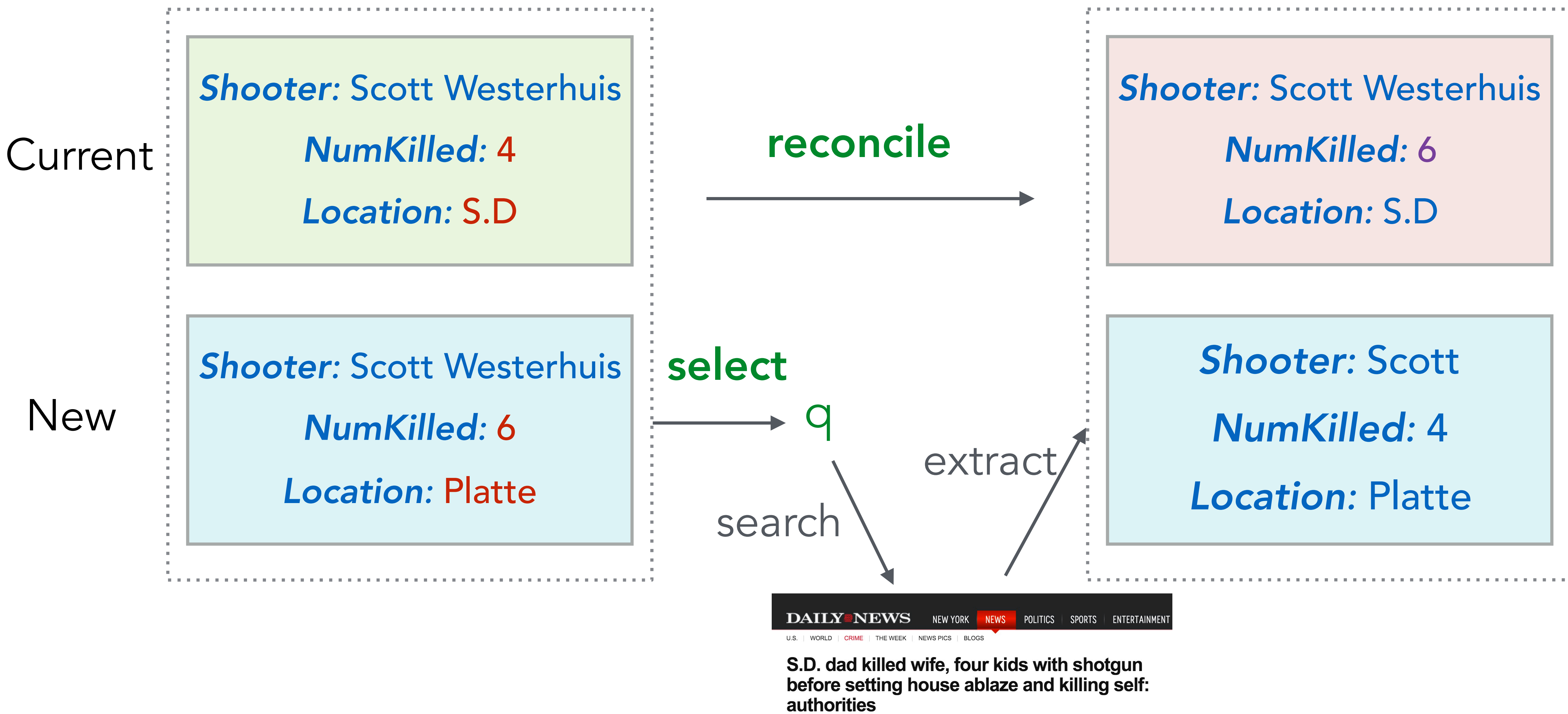
RL: Actions



1. **Reconcile (d)** current values and new values.
   a) Pick a single value, all values or no value from new set
   b) Stop

$\boxed{a = (d, q)}$

RL: Actions



2. Select next **query (q).**

$\boxed{a = (d, q)}$

Queries

Query templates are induced **automatically**

- Title of original article
- Content words having high mutual information with gold values

<title>
<title> + (suspect | shooter | said | men | arrested | …)
<title> + (injured | wounded | victims | shot | …)

Rewards

- Change in accuracy

Current Values

✓ ***Shooter:*** Scott Westerhuis
✓ ***NumKilled:*** 6
✓ ***NumWounded:*** 0
✓ ***Location:*** Platte

Previous Values

✓ ***Shooter:*** Scott Westerhuis
✓ ***NumKilled:*** 6
✗ ***NumWounded:*** 1
✓ ***Location:*** Platte

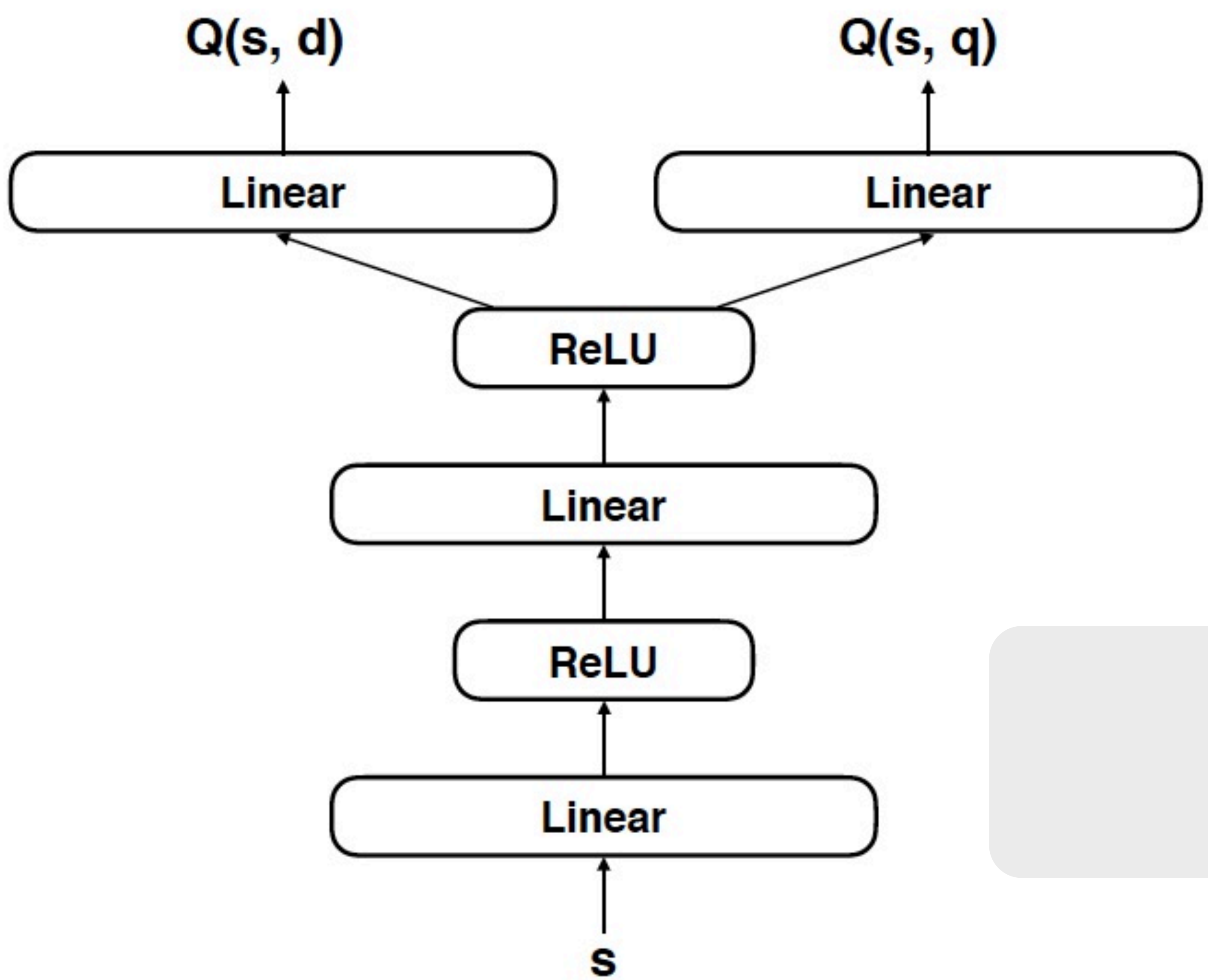
R = 1

$$R(s,a) = \sum_{\text{entity}\, j} \text{Acc}(e^j_{cur}) - \text{Acc}(e^j_{prev})$$

- Small penalty for each transition

Deep Q-Network

State space is continuous: requires function approximation



$$Q(s, a) \approx Q(s, a; \theta)$$

Trained to maximize cumulative reward

Experiments

Two domains with multiple information sources:

1. Mass shootings in the United States
2. Adulteration events from Foodshield EMA

Number	Shootings			Adulteration		
	Train	**Test**	**Dev**	**Train**	**Test**	**Dev**
Source articles	306	292	66	292	148	42
Downloaded articles	8201	7904	1628	7686	5333	1537

Base Extraction Model

killed

wounded

One person was killed and four other people were wounded in a shooting in Chicago's West Englewood neighborhood Monday night. The youngest victim is 13 years old.

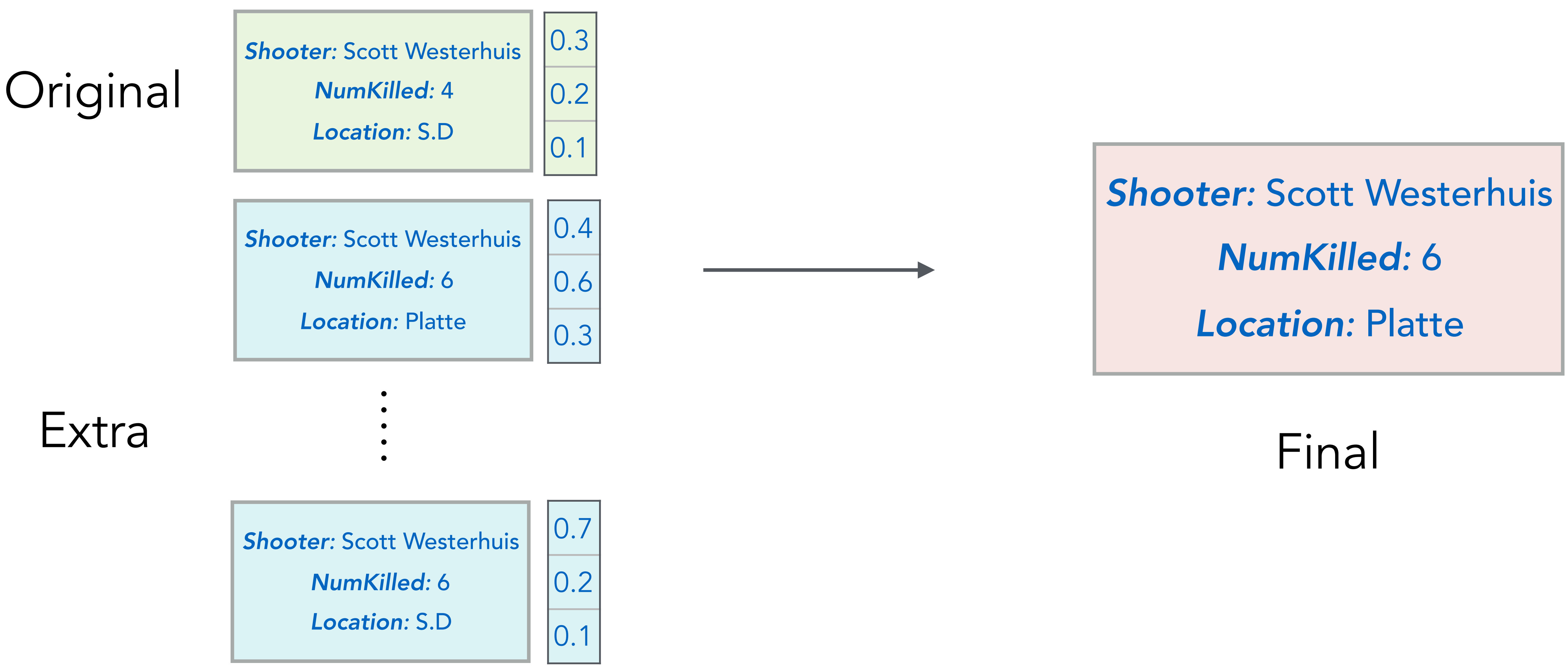
location

Maximum Entropy classifier with contextual features to label each word

Baselines (1)

Simple Reconciliation systems:

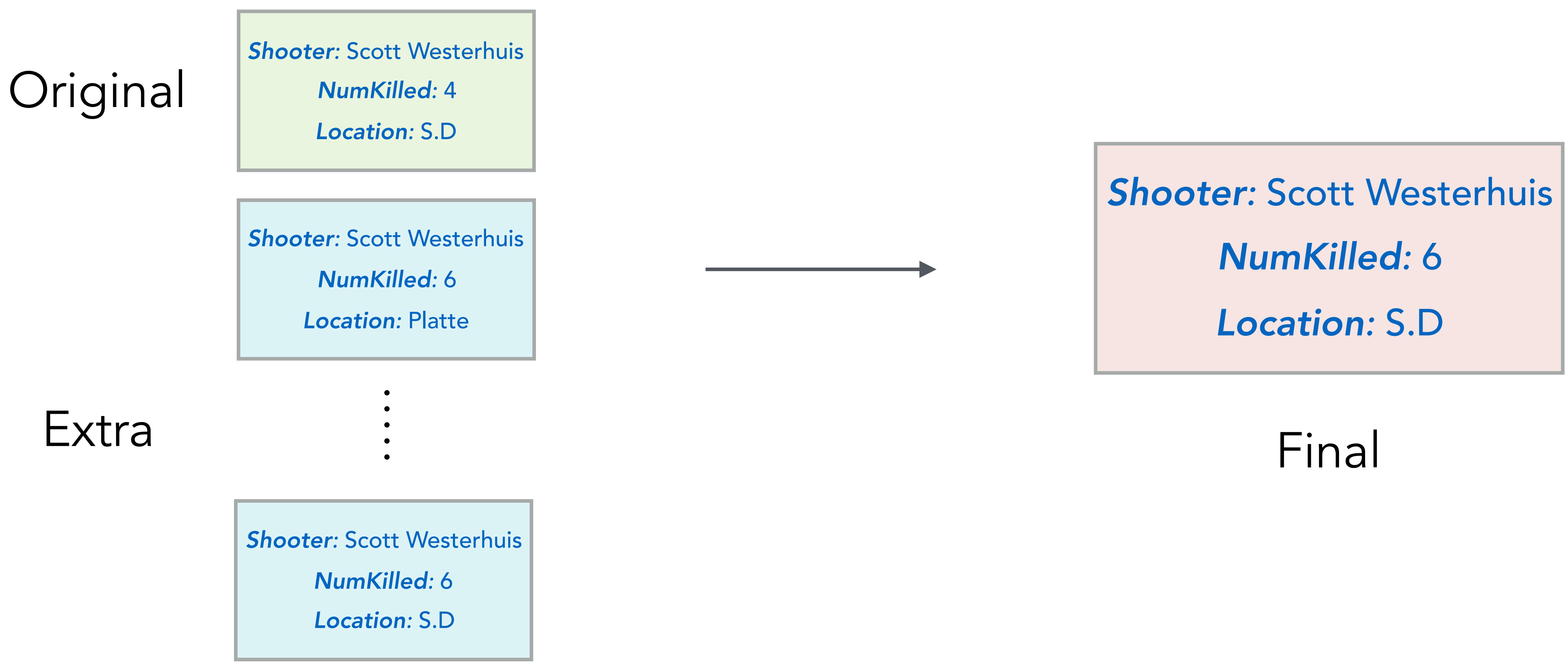
- *Confidence-based*: Choose entity value with highest confidence



Baselines (1)

Simple Reconciliation systems:

- *Majority-based*: Choose entity value extracted the most from all articles on the event



Baselines (2)

Meta-classifier for reconciliation:

- Operates over same input space S and same set of reconciliation decisions as the RL agent.

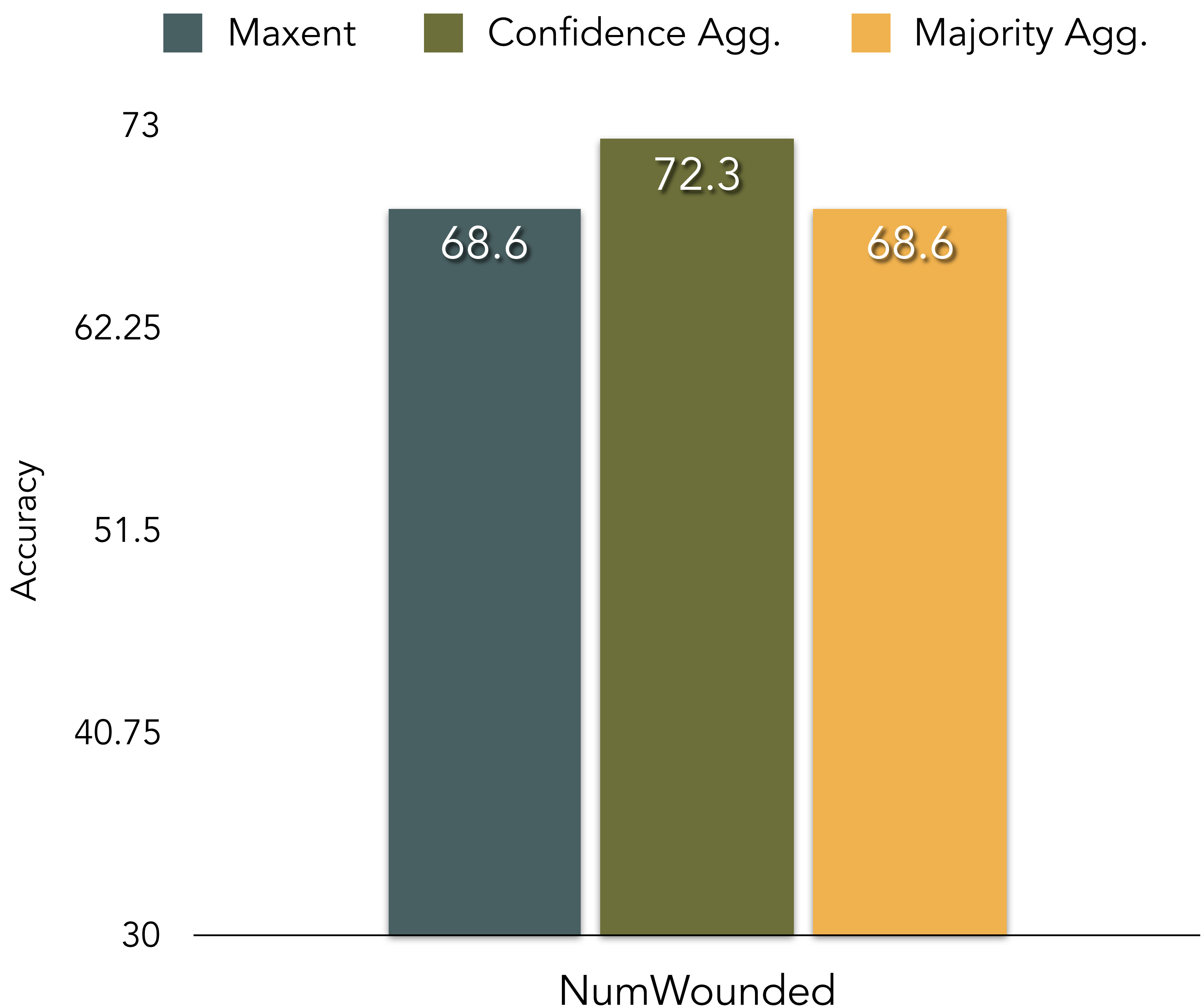
< 0.3, 0.2, 0.1, 0.4, 0.6, 0.3, 1, 0, 0, 0.65, 0, 0, 1, … , 0, 0 >

currentConf newConf matches docSim context

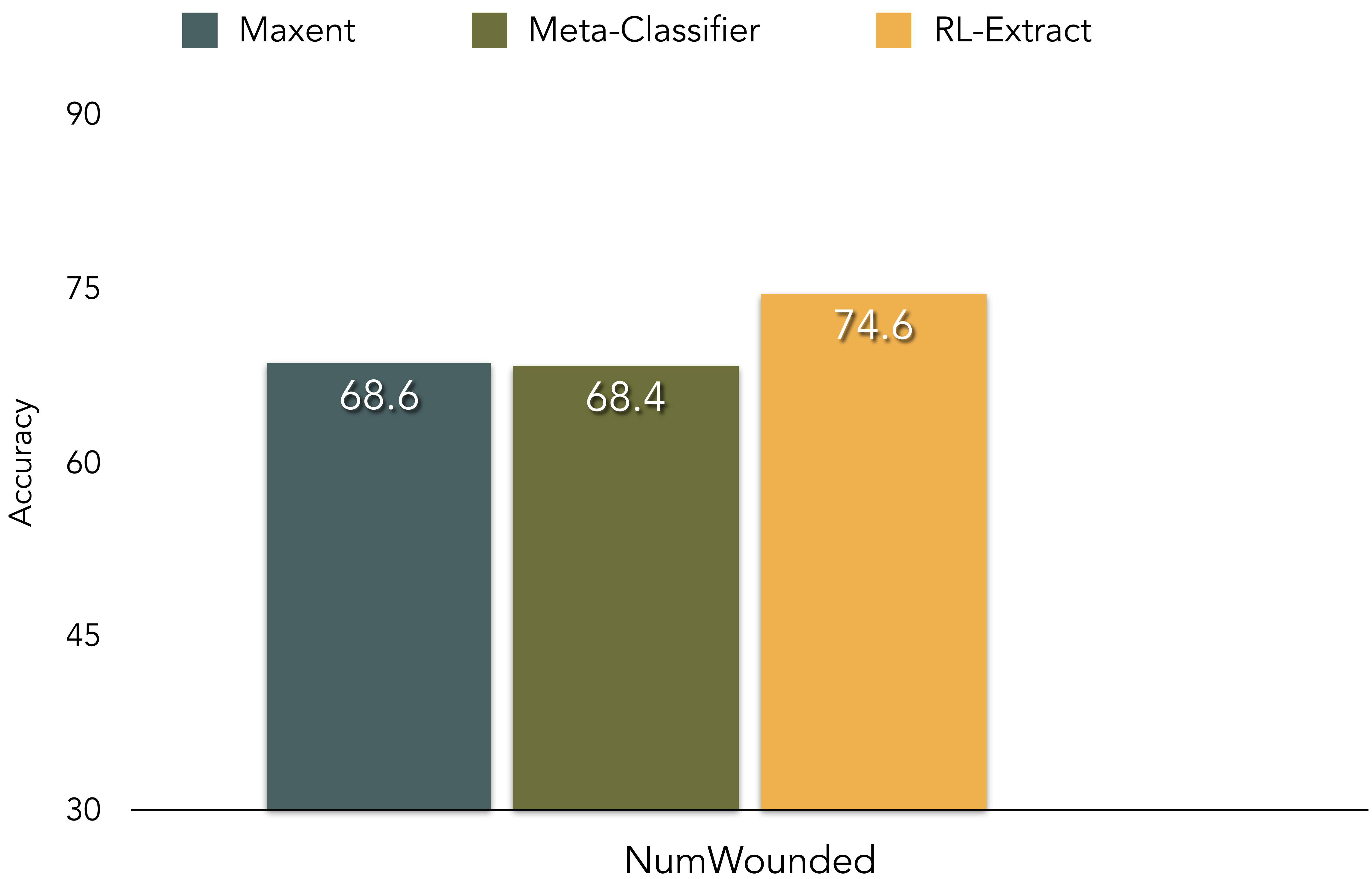
- Each state has values from original article and one extra article

- Applied to all extra articles followed by confidence-based aggregation

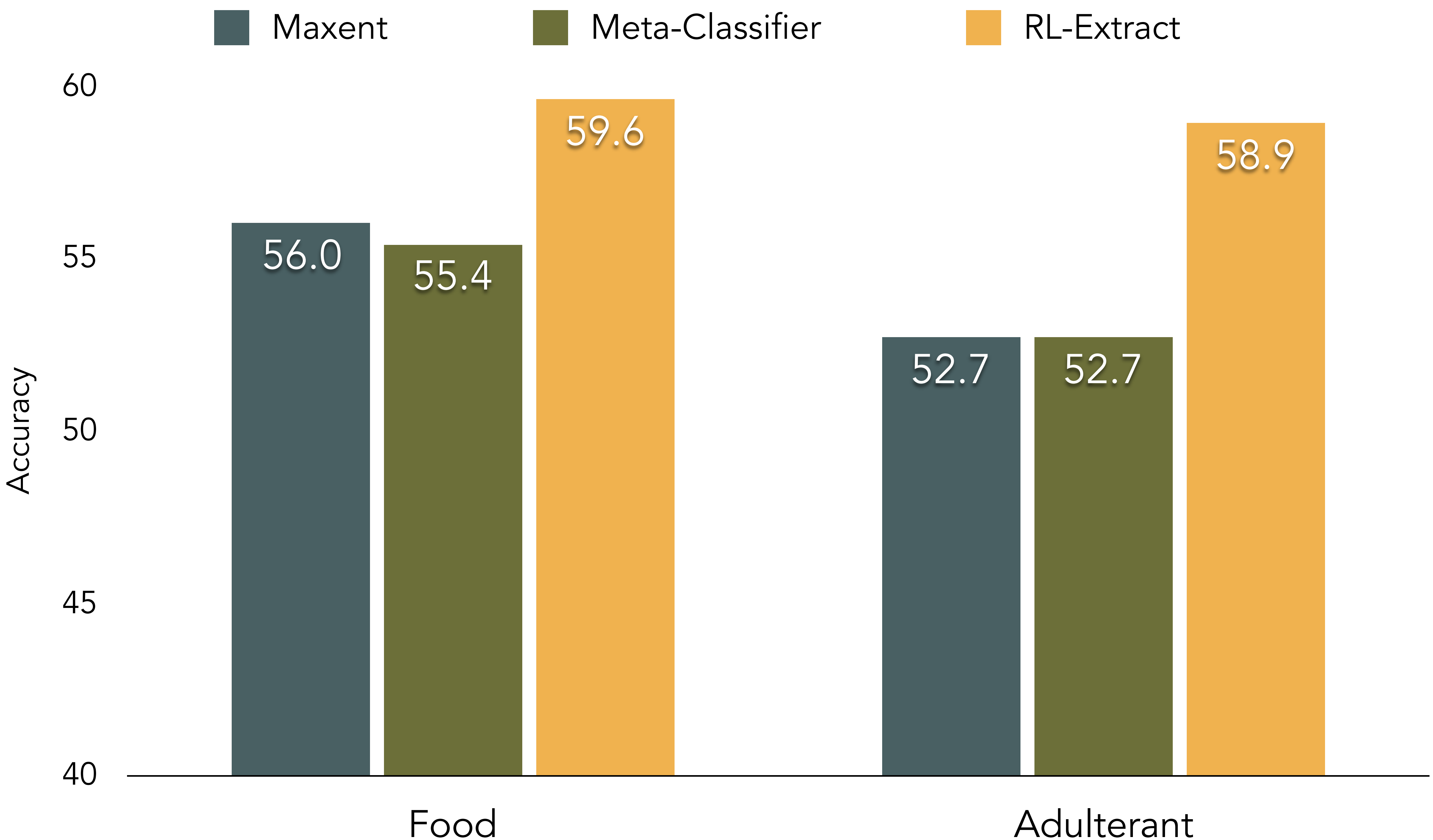
Accuracy (Shootings)



Accuracy (Shootings)



Accuracy (Adulterations)



Examples

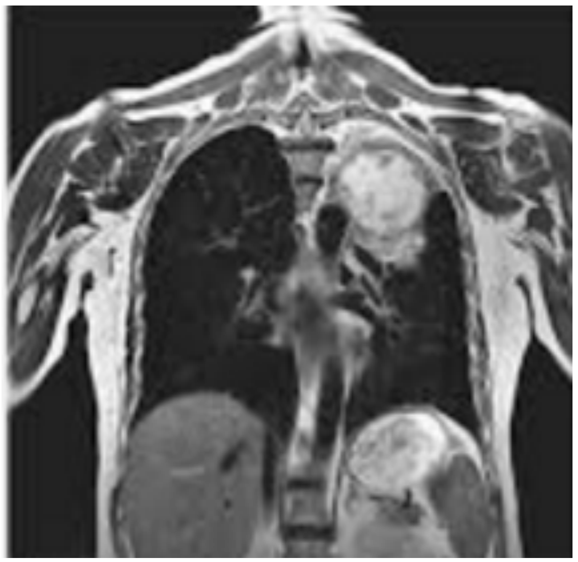
	Text	Shooter Name
Basic Extractor	A source tells Channel 2 Action News that Thomas Lee has been arrested in Mississippi … Sgt . Stewart Smith, with the Troup County Sheriff's office, said.	Stewart
RL-Extract	Lee is accused of killing his wife, Christie; …	Lee

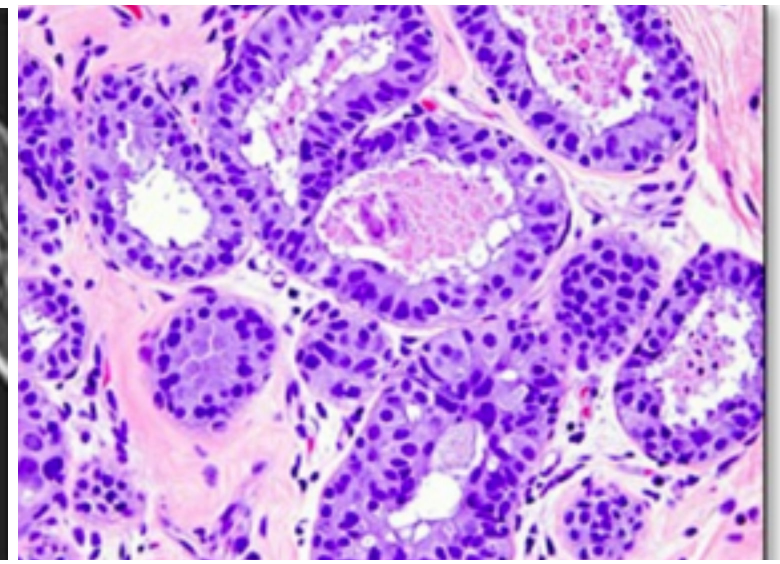
Examples

	Text	# Killed
Basic Extractor	Shooting leaves 25 year old Pittsfield man dead , 4 injured	0
RL-Extract	One man is dead after a shooting Saturday night at the intersection of Dewey Avenue and Linden Street.	1

Our system finds alternative sources of information for reliable extraction

Predicting Clinical Outcomes from EHR





AST	8~ 40 U/L
ALT	5~ 40 U/L
LDH	50~400 U/L
ALP	80~280 U/L
γ-GTP	0~ 50 U/L
ZTT	4~ 12 KU
TTT	0~ 5 KU
Total bilirubin	0.4~0.9 mg/dℓ
Direct bilirubin	0~0.4 mg/dℓ
Indirect bilirubin	0.4~0.5 mg/dℓ
Total protein	6.0~9.0 g/dℓ

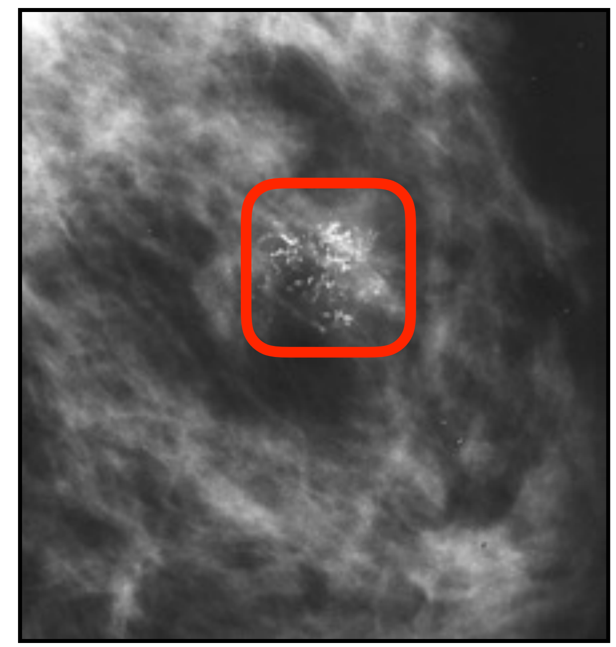
Clinical Prediction

Doctors need to know reasons behind predictions!

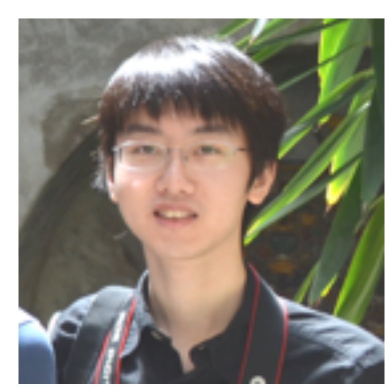
Interpretable Neural Models

- Complex (neural) models come at the cost of interpretability
- Applications often need interpretable justifications — rationales.

There is no evidence of extranodal extension.
BREAST (RIGHT), EXCISIONAL BIOPSY:
INVASIVE DUCTAL CARCINOMA (SEE TABLE #1). DUCTAL CARCINOMA IN-SITU, GRADE 1. ATYPICAL DUCTAL HYPERPLASIA. LOBULAR NEOPLASIA (ATYPICAL LOBULAR HYPERPLASIA). TABLE OF PATHOLOGICAL FINDINGS #1 INVASIVE CARCINOMA
… …



prediction: high risk



Tao Lei

Tommi Jaakkola

Motivation

- Complex (neural) models come at the cost of interpretability
- Applications often need interpretable justifications — rationales.

this beer pours ridiculously clear with tons of carbonation that forms a rather impressive rocky head that settles slowly into a fairly dense layer of foam. this is a real good lookin' beer, unfortunately it gets worse from here ... first, the aroma is kind of bubblegum-like and grainy. next, the taste is sweet and grainy with an unpleasant bitterness in the finish. ... ... overall, the fat weasel is good for a fairly cheap buzz, but only if you like your beer grainy and bitter .

Ratings

Look: 5 stars

Aroma: 2 stars

review with rationales

Our goal: learn to extract rationales behind predictions

Problem Setup

Interpretability via providing concise evidence from input

Rationales (evidence) should be:

- short and coherent pieces
- sufficient for correct prediction

Rationales are not provided during training

in contrast to *(Zaidan et al., 2007; Marshall et al.,2015; Zhang et al., 2016)*

Use powerful neural nets to avoid accuracy loss

in contrast to *(Thrun, 1995; Craven and Shavlik, 1996; Ribeiro et al., 2016)*

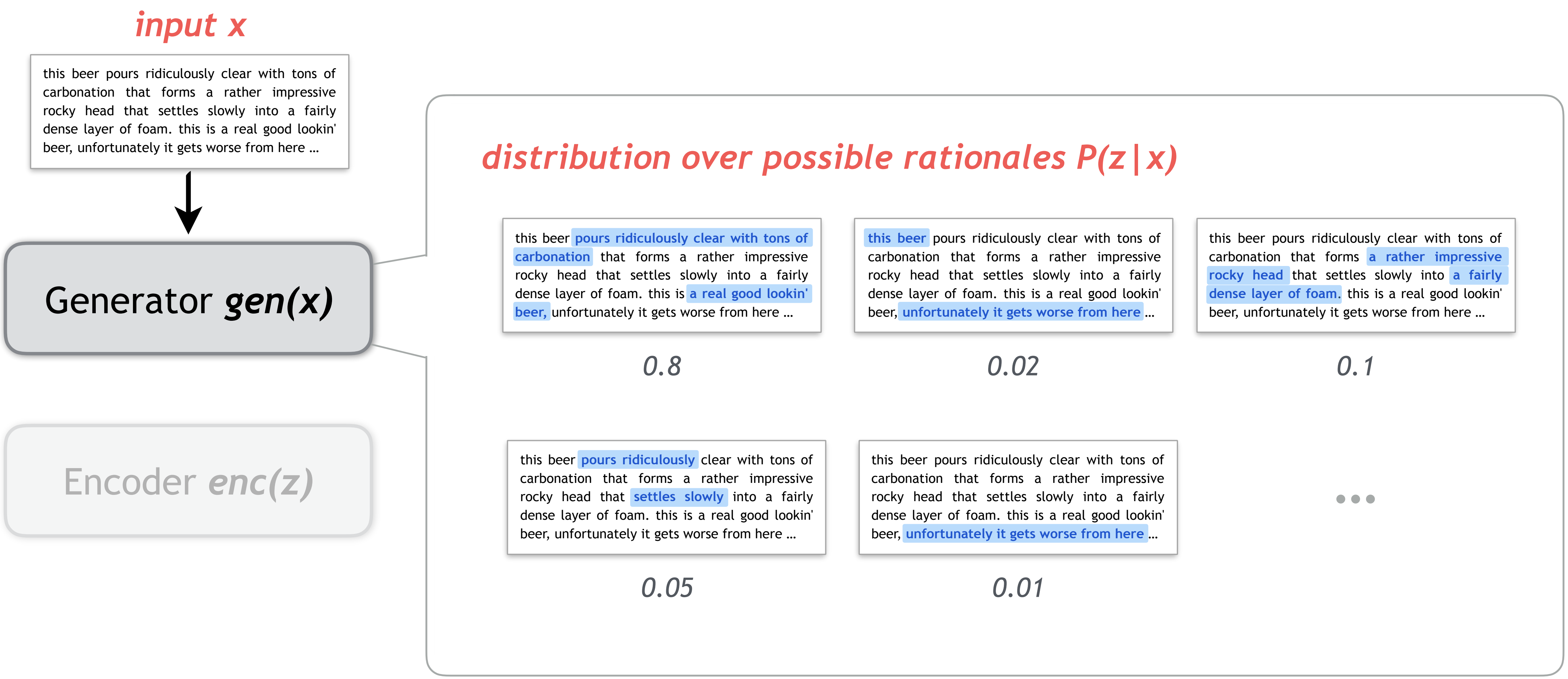
Model Architecture

Generator ***gen(x)***

Encoder ***enc(z)***

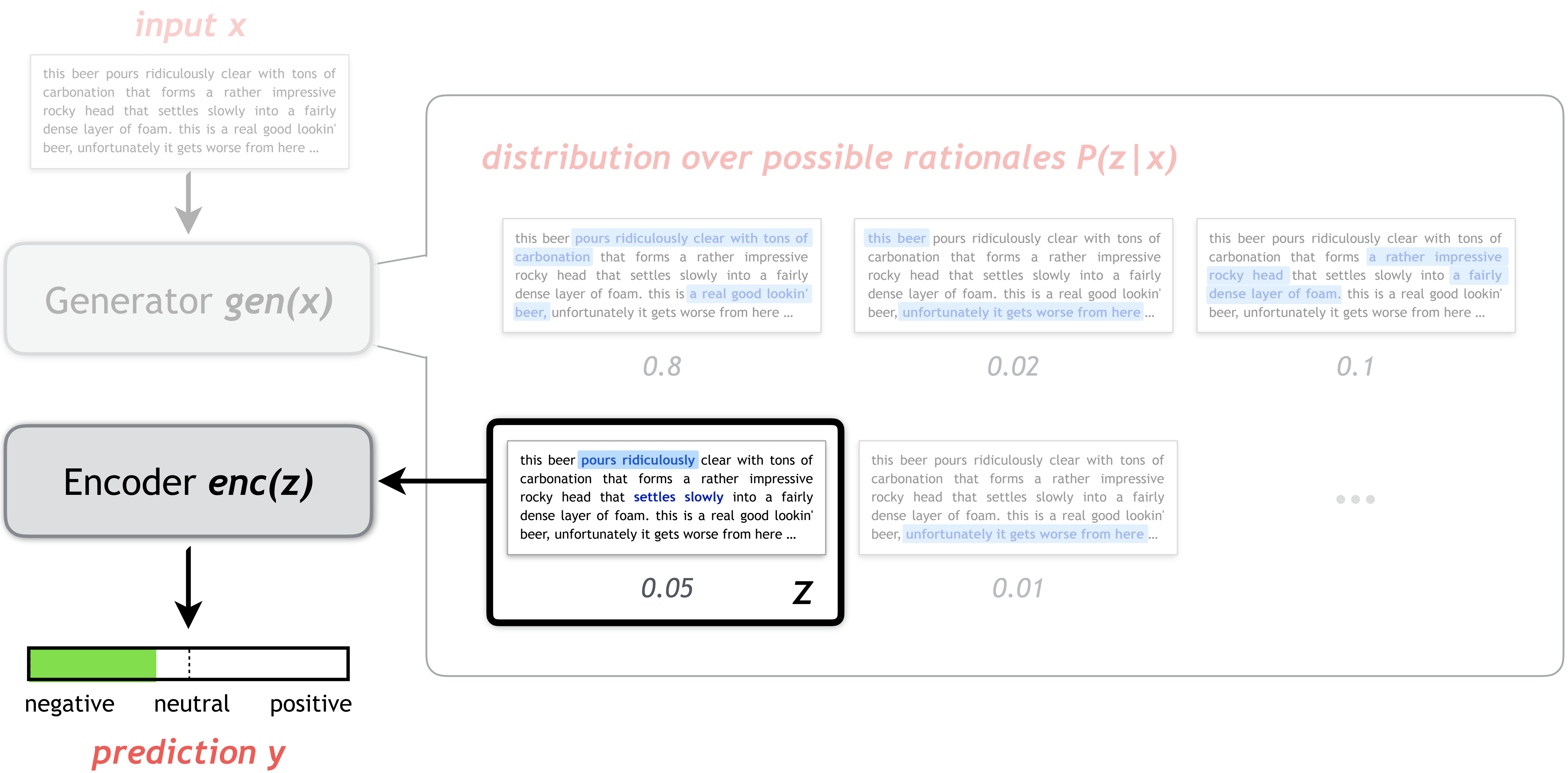
two modular components ***gen()*** and ***enc()***

Model Architecture



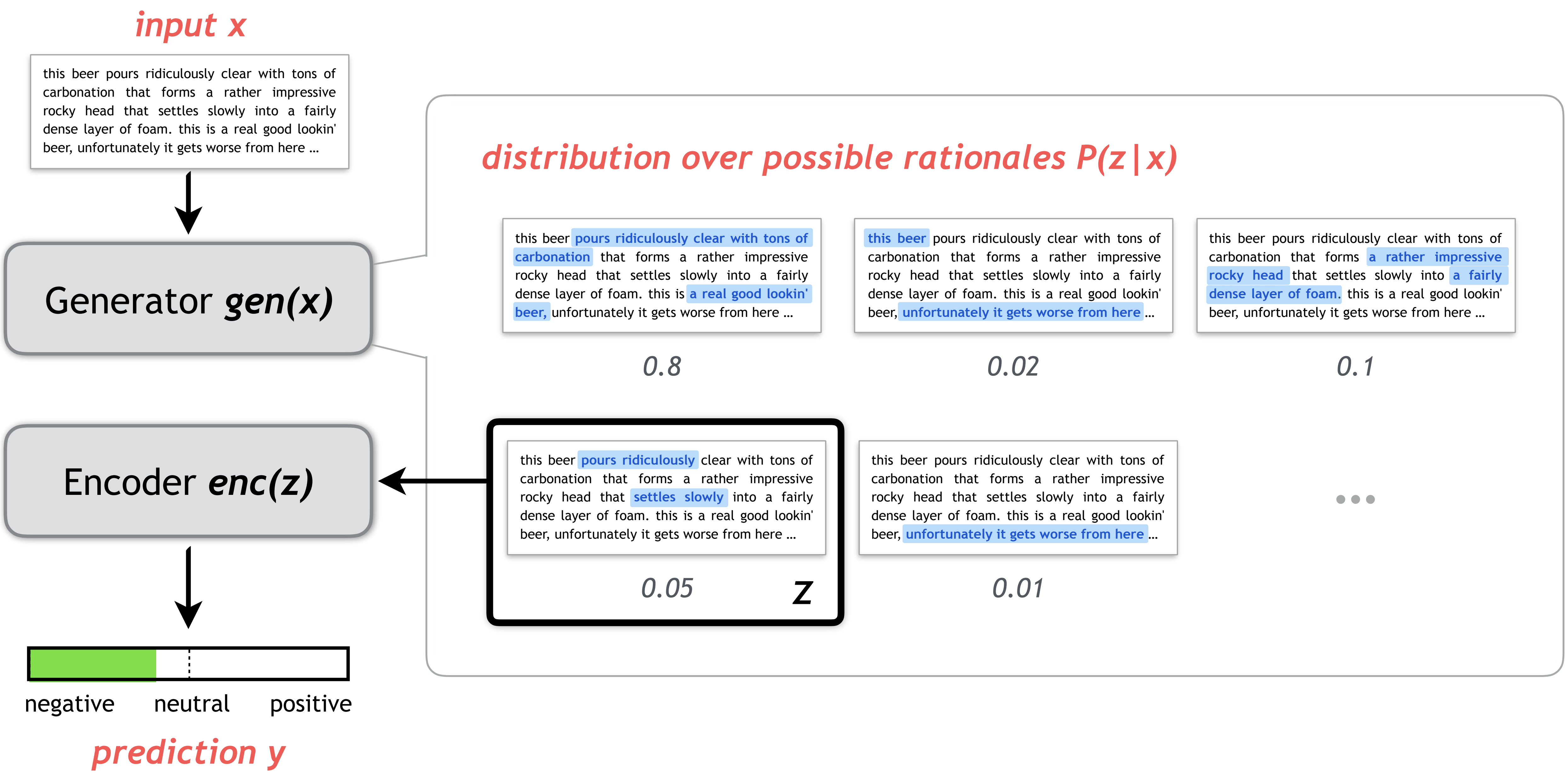
generator specifies the distribution of rationales

Model Architecture



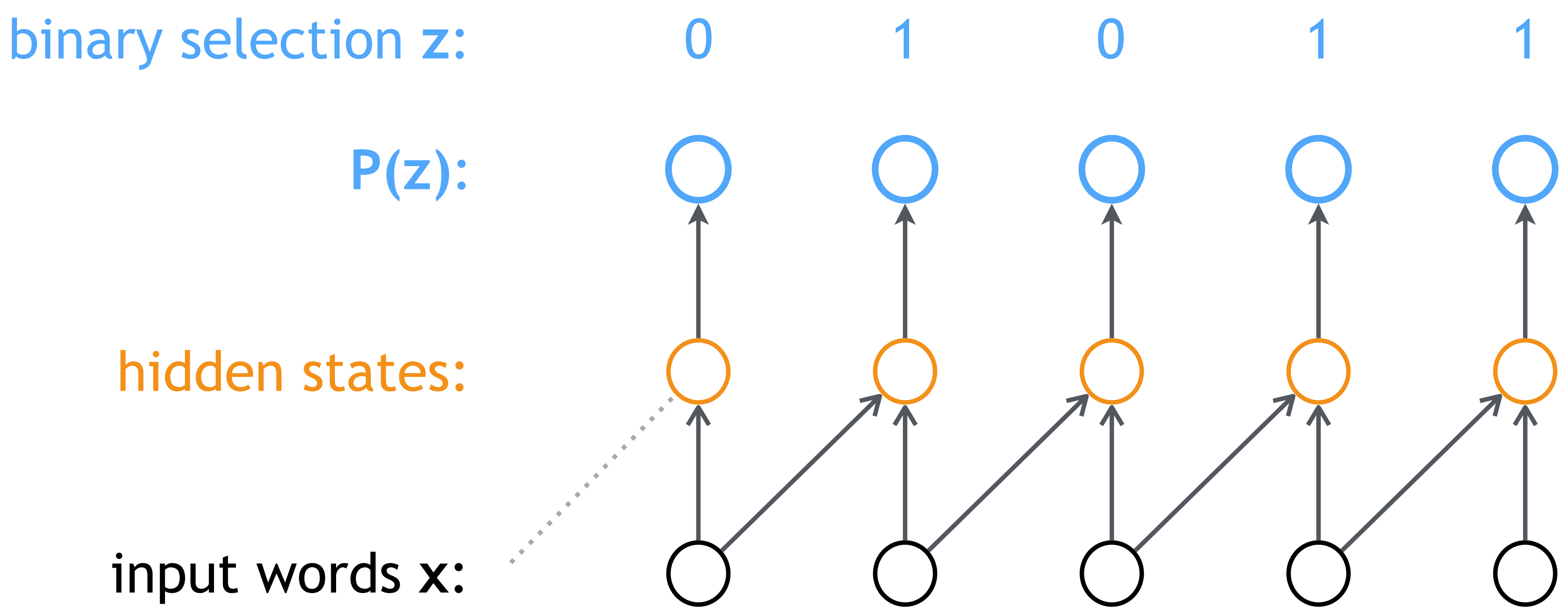
encoder makes prediction given rationale

Model Architecture



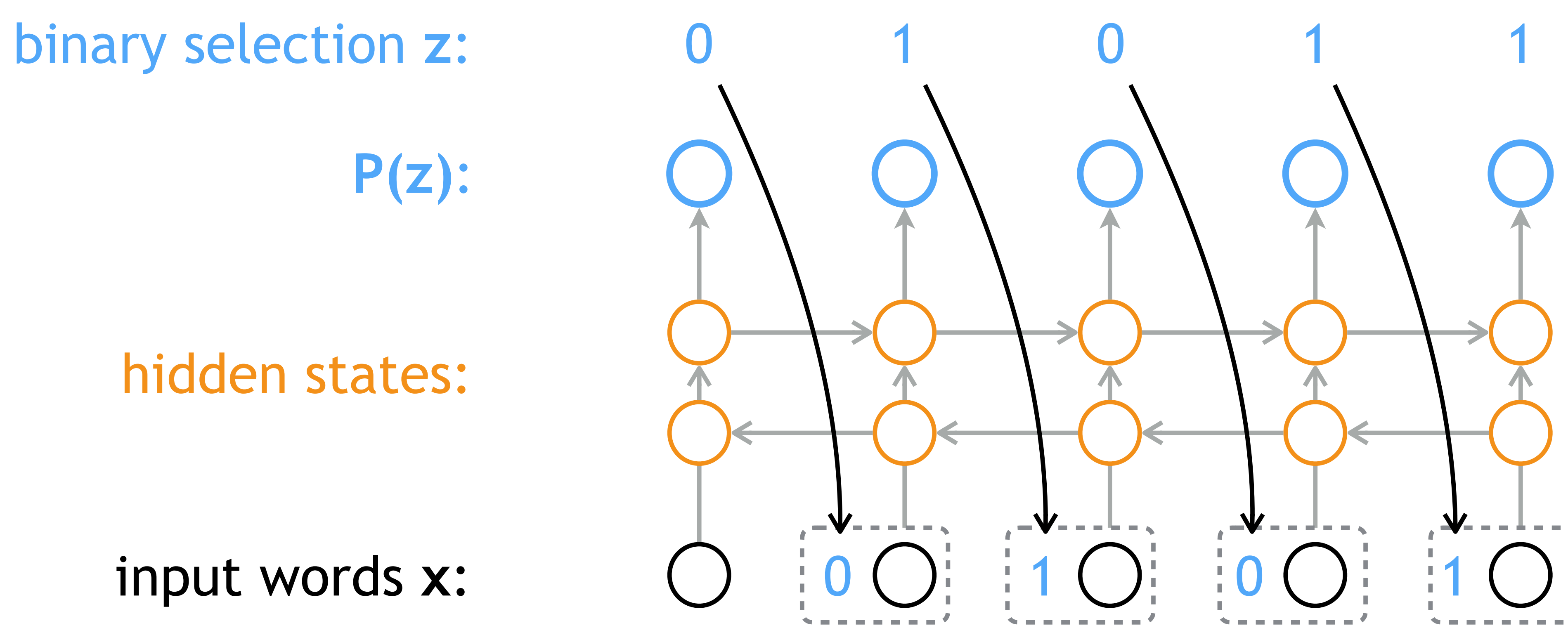
two components optimized jointly

Generator Implementations



independent selection, feedforward net

Generator Implementations



dependent selection, bi-directional RNNs

choose networks based on the data/application

Training Objective

$$\mathrm{cost}(\mathbf{z},\mathbf{y}) = \underbrace{\mathrm{loss}(\mathbf{z},\mathbf{y})}_{\textit{sufficiency}} + \underbrace{\lambda_1 |\mathbf{z}|_1}_{\textit{sparsity}} + \underbrace{\lambda_2 \sum_i |\mathbf{z}_i - \mathbf{z}_{i-1}|}_{\textit{coherency}}$$

sufficiency *correct prediction*

sparsity *rationale is short*

coherency *continuous selection*

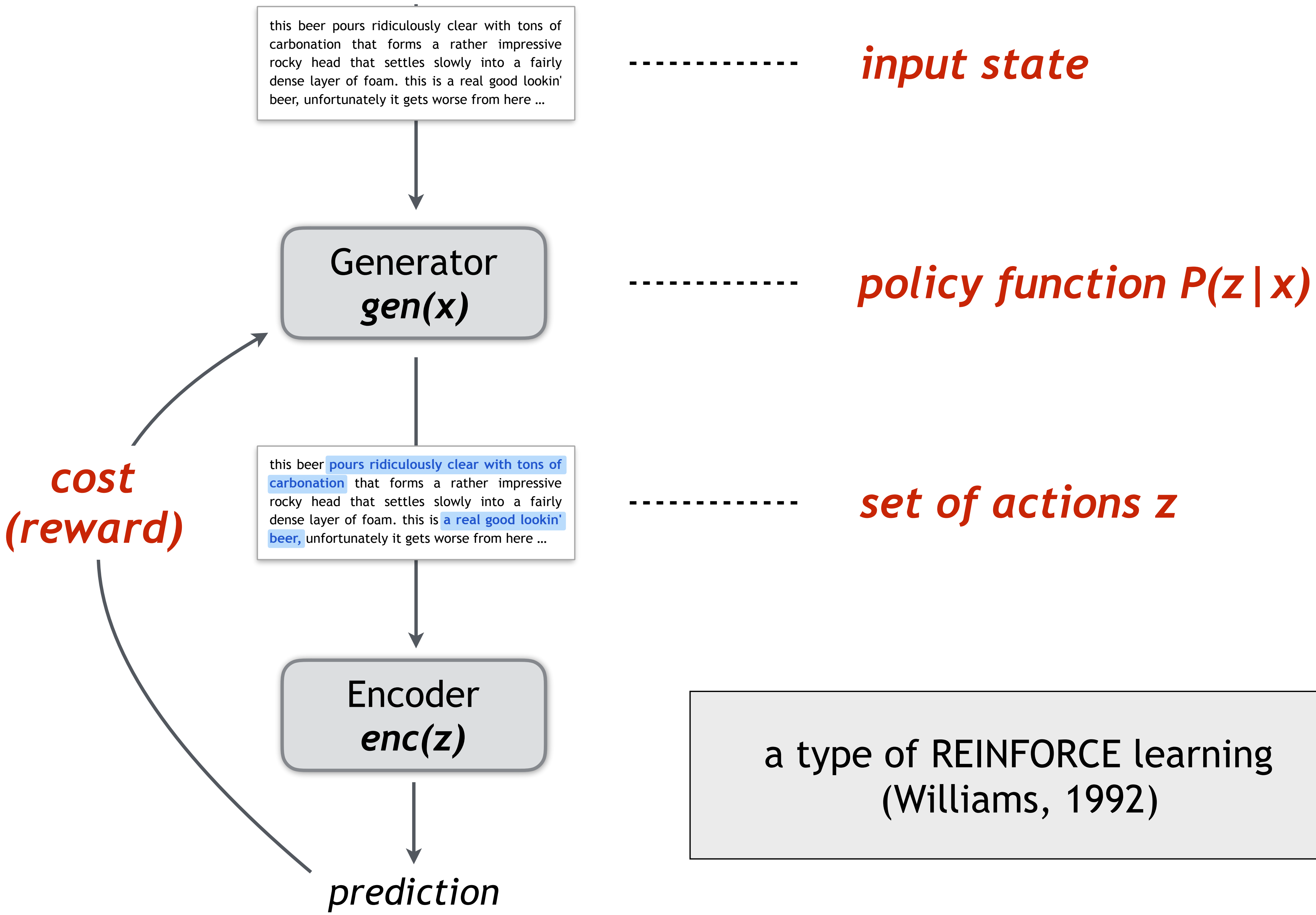
- receive this training signal after z is produced

Minimizing expected cost:

$$\min_{\theta} \sum_{(\mathbf{x},\mathbf{y})\in D} \mathbb{E}_{\mathbf{z}\sim \mathrm{gen}(\mathbf{x})} \left[\mathrm{cost}(\mathbf{z},\mathbf{y}) \right]$$

- intractable because summation over z is exponential

Learning as Policy Gradient Method



a type of REINFORCE learning
(Williams, 1992)

Experiments

Three real-world datasets and applications for evaluation:

Predicting sentiment for product reviews

Parsing medical pathology reports

Finding similar posts on QA forum

Evaluation: Product Review

Dataset: multi-aspect beer reviews from *BeerAdvocate* (McAuley et al, 2012) 1.5m in total

1,000 reviews annotated at sentence level with aspect label (used only for evaluation)

Task: predict ratings and rationales for each aspect

this beer **pours ridiculously clear with tons of carbonation** that forms a rather impressive rocky head that settles slowly into a fairly dense layer of foam. **this is a real good lookin' beer**, unfortunately it gets worse from here … first, **the aroma is kind of bubblegum-like and grainy.** next, the taste is sweet and grainy with an unpleasant bitterness in the finish. … … overall, the fat weasel is good for a fairly cheap buzz, but only if you like your beer grainy and bitter .

Ratings

Look: 5 stars

Aroma: 2 stars

Evaluation: Product Review

Set-up: ratings are fractional; treat the task as **regression** following (McAuley et al, 2012)

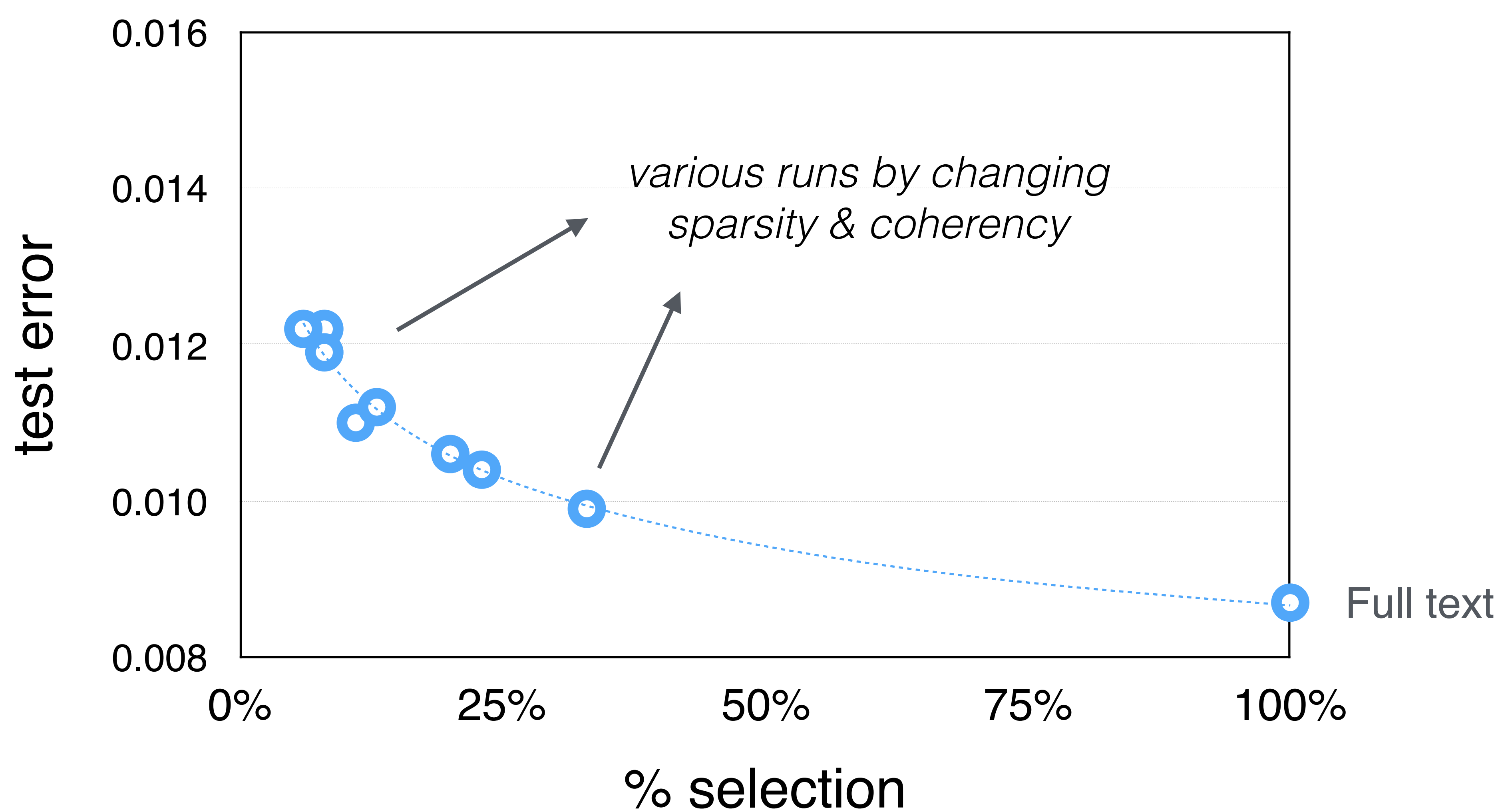
use recurrent networks for ***gen()*** and ***enc()***

Metrics: **precision**:
percentage of selected words in correct sentences

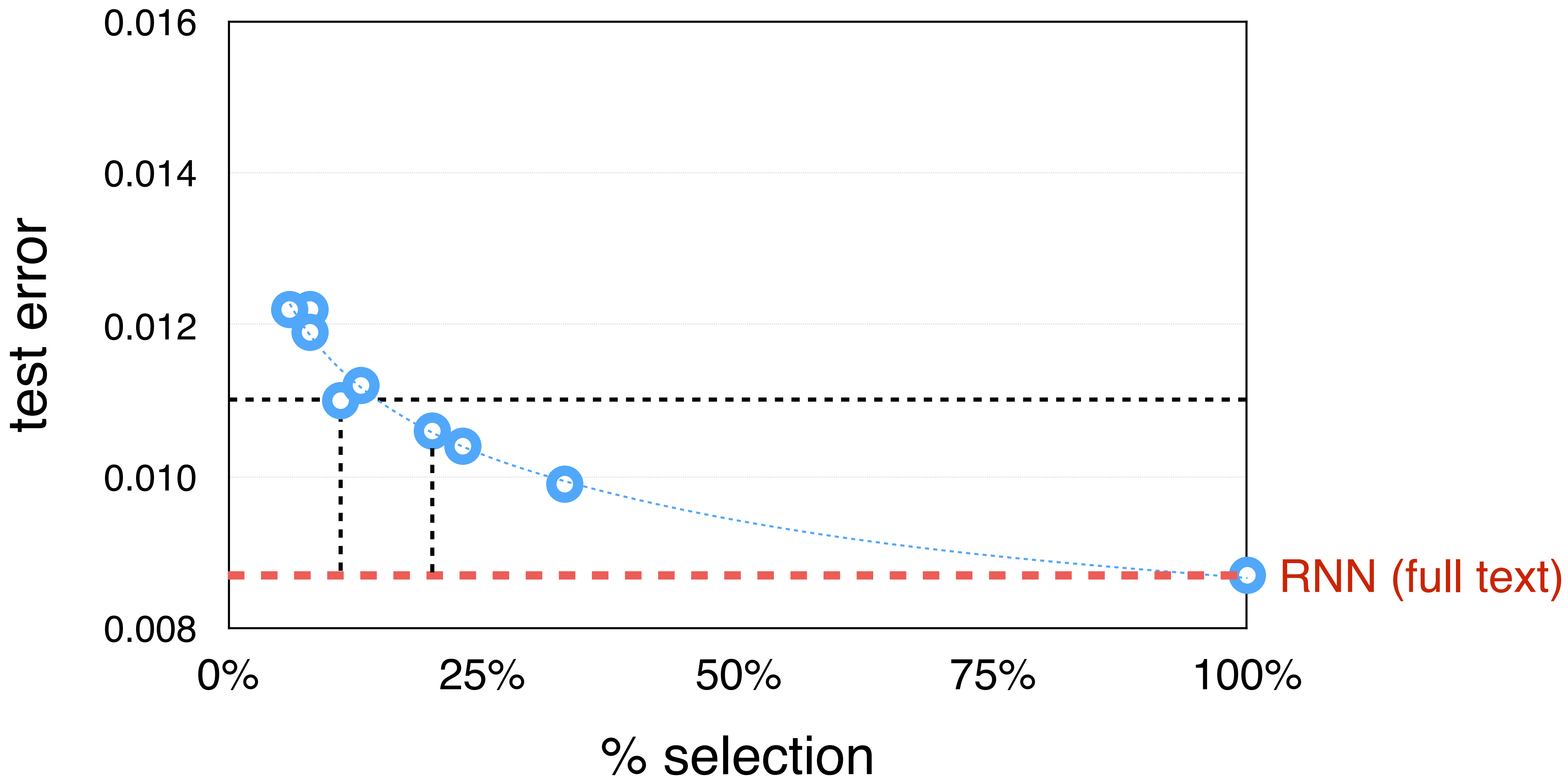
mean squared error on sentiment prediction

Baselines: SVM classifier
attention-based RNN

Sentiment Prediction

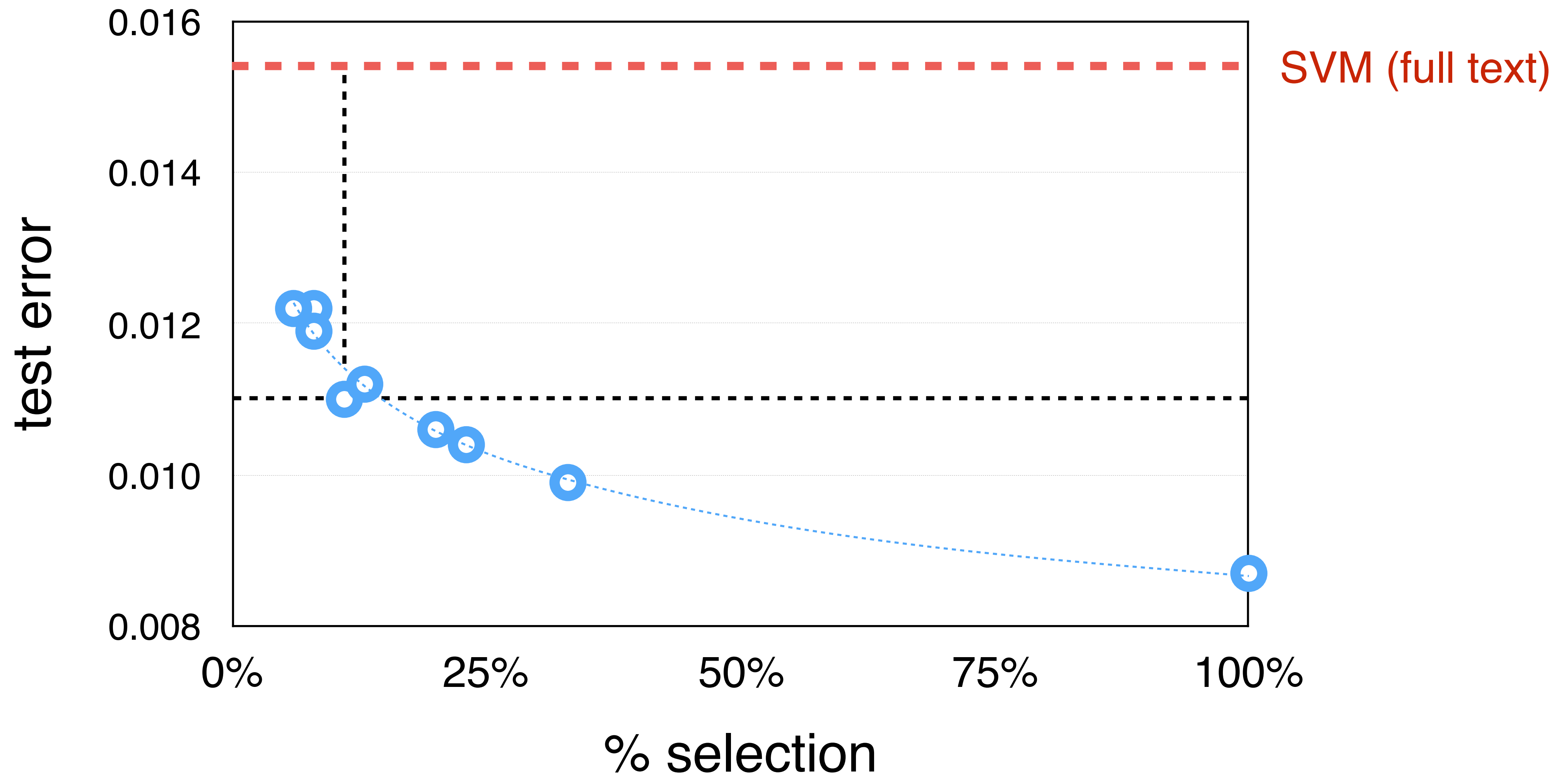


Sentiment Prediction



rationales getting close performance to full text

Sentiment Prediction



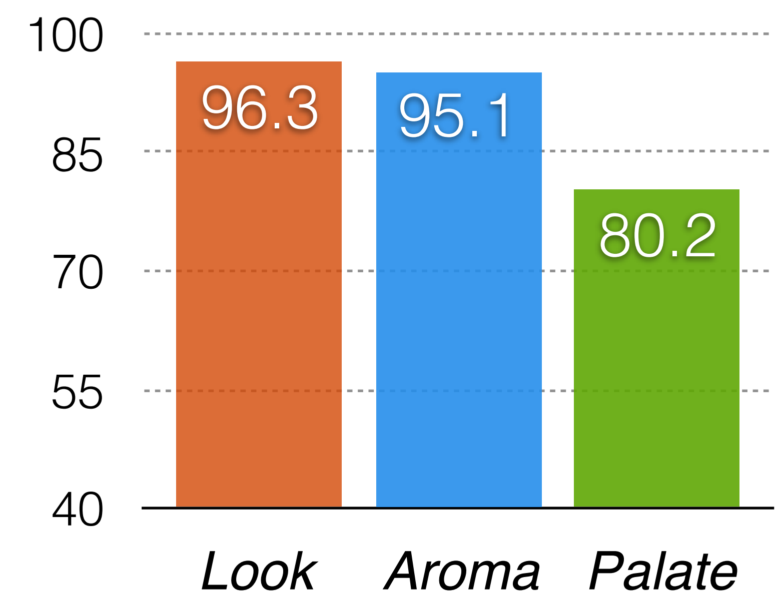
advantage of neural models over linear classifiers still clear

Precision of Rationales

Examples and precisions of rationales

a beer that is not sold in my neck of the woods , but managed to get while on a roadtrip . poured into an imperial pint glass with **a generous head that sustained life throughout** . nothing out of the ordinary here , but a good brew still . body **was kind of heavy , but not thick** . the **hop smell was excellent and enticing . very drinkable**

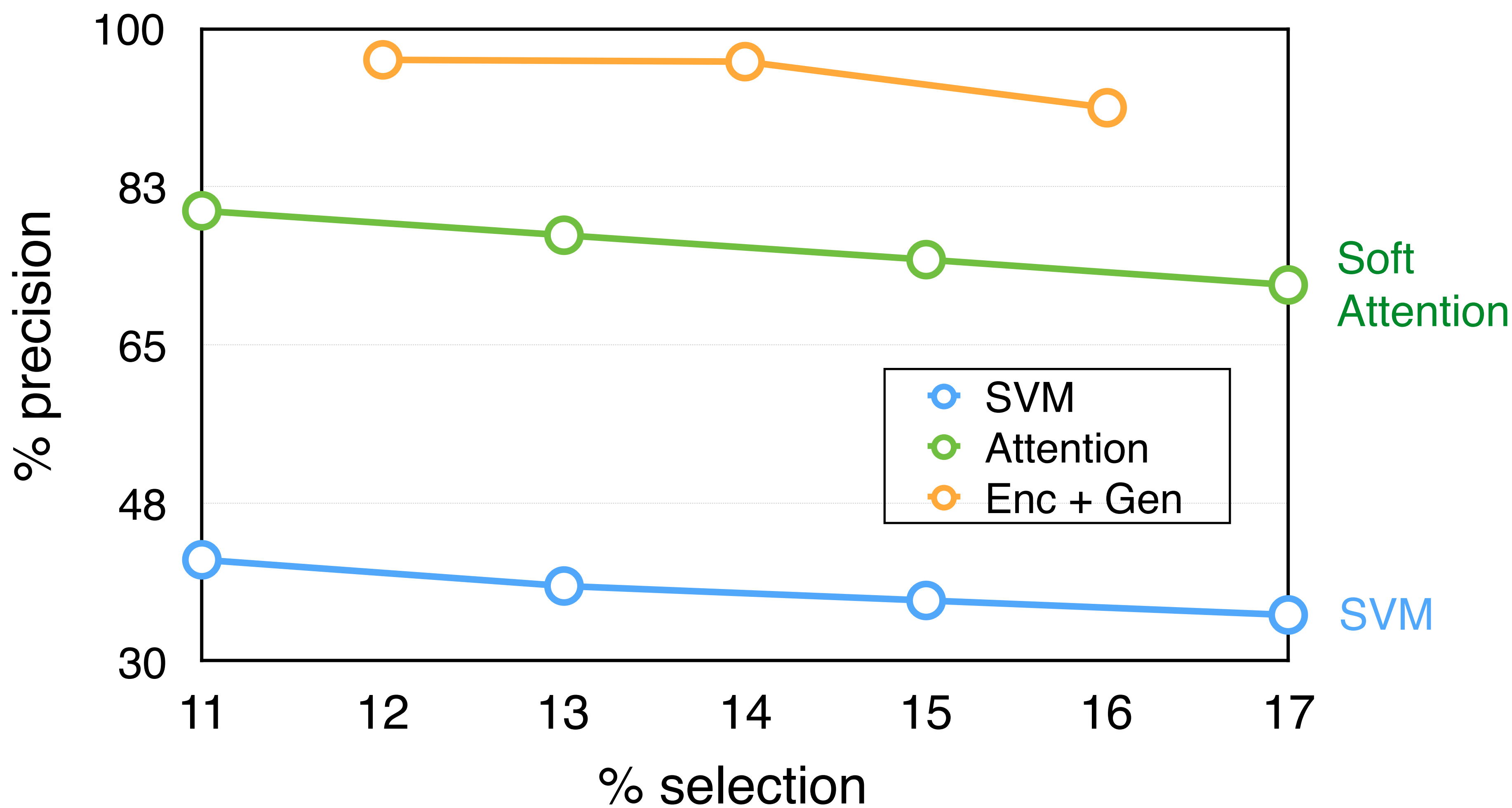
poured into a snifter . **produces a small coffee head that reduces quickly . black as night** . pretty typical imp . **roasted malts** hit on the nose . **a little sweet chocolate follows** . big toasty character on the taste . in between i 'm getting plenty of dark chocolate and some bitter espresso . it finishes with hop bitterness . **nice smooth mouthfeel with perfect carbonation for the** style . overall a nice stout i would love to have again , maybe with some age on it .



more examples available at

https://github.com/taolei87/rcnn/tree/master/code/rationale

Precision of Rationales



proper modeling leads to better rationale

Evaluation: Parsing Pathology Report

Dataset: patients' pathology reports from hospitals such as MGH

Task: check if a disease/symptom is positive in text

binary classification for each category

Statistics: several thousand report for each category

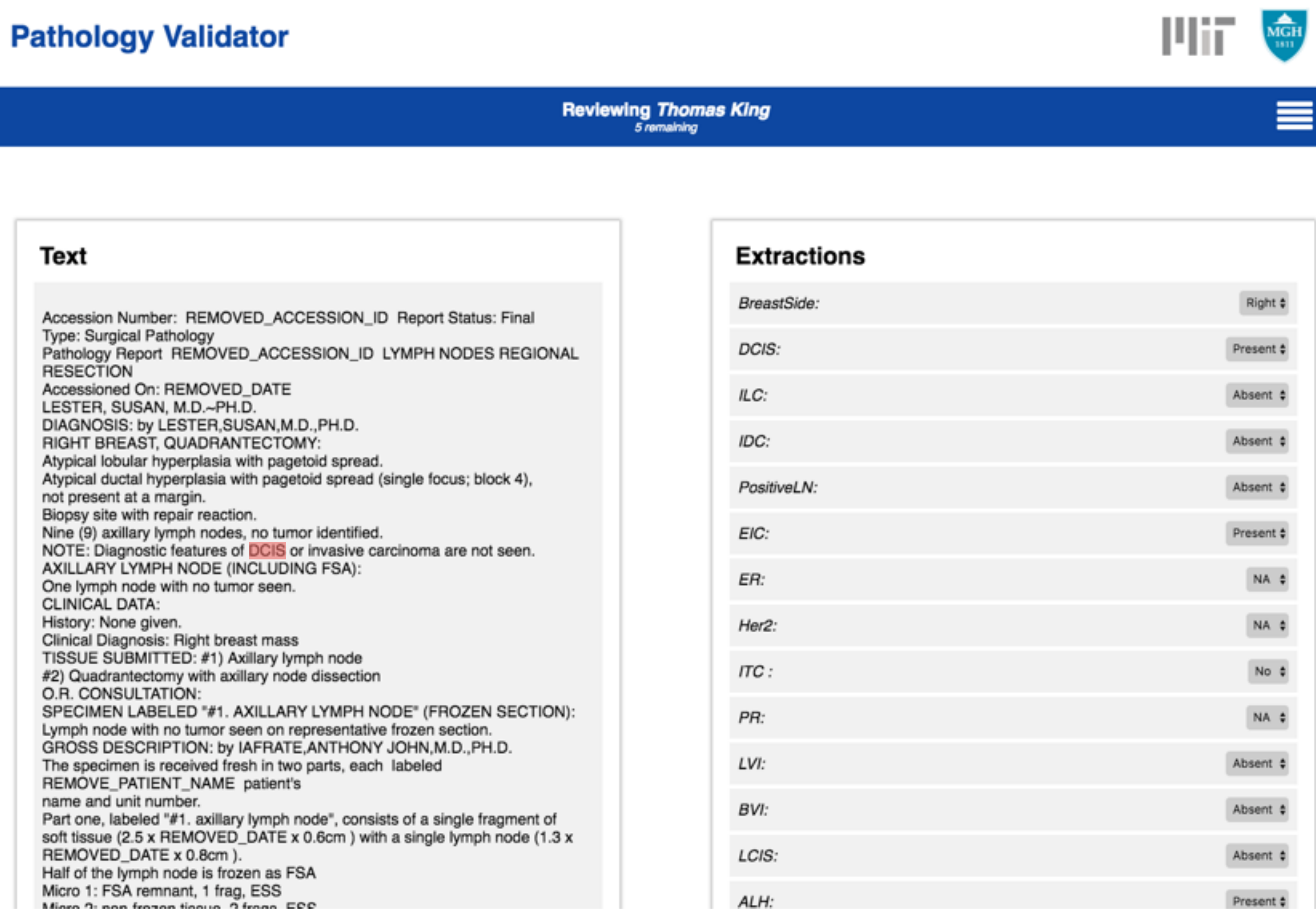
pathology report is long (>1000 words) but structured

Model: use CNNs fro ***gen()*** and ***enc()***

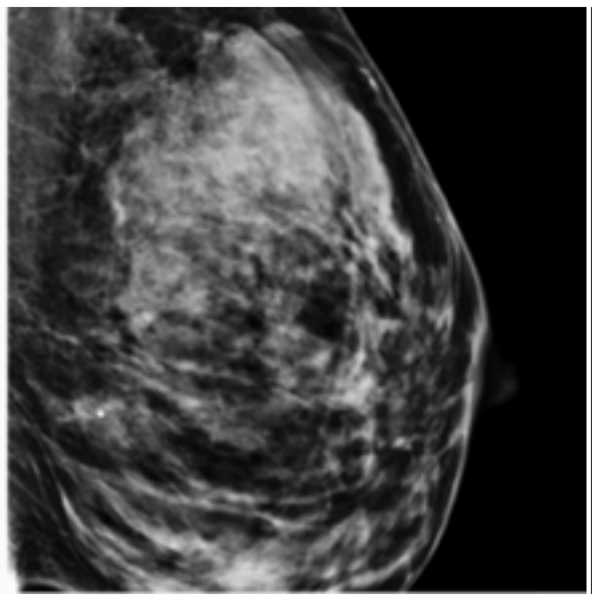
Evaluation: Parsing Pathology Report

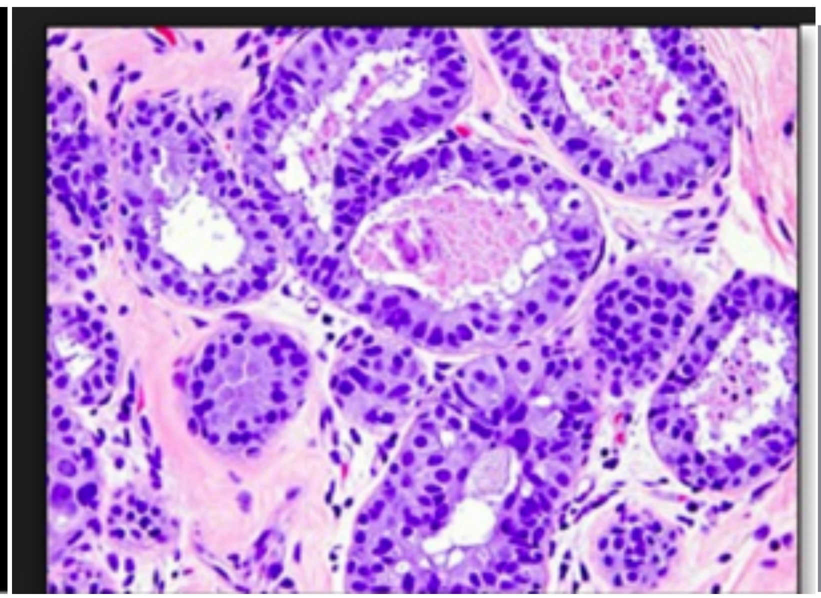
Category:		*F-score:*
IDC	*Accession Number* <unk> *Report Status* Final *Type* Surgical Pathology … *Pathology Report:* LEFT BREAST ULTRASOUND GUIDED CORE NEEDLE BIOPSIES … **INVASIVE DUCTAL CARCINOMA poorly differentiated modified Bloom** Richardson grade III III measuring at least 0 7cm in this limited specimen Central hyalinization is present within the tumor mass but no necrosis is noted No lymphovascular invasion is identified No in situ carcinoma is present Special studies were performed at an outside institution with the following results not reviewed ESTROGEN RECEPTOR NEGATIVE PROGESTERONE RECEPTOR NEGATIVE …	98%
LCIS	… **Extensive** LCIS DCIS **Invasive** carcinoma of left breast FINAL DIAGNOSIS BREAST **LEFT LOBULAR CARCINOMA IN SITU PRESENT** ADJACENT TO PREVIOUS BIOPSY SITE SEE NOTE CHRONIC INFLAMMATION ORGANIZING HEMORRHAGE AND FAT NECROSIS BIOPSY SITE NOTE There is a second area of focal lobular carcinoma in situ noted with pagetoid spread into ducts No vascular invasion is seen The margins are free of tumor No tumor seen in 14 lymph nodes examined BREAST left breast is a <unk> gram 25 x 28 x 6cm left …	97%
LVI	FINAL DIAGNOSIS BREAST RIGHT EXCISIONAL BIOPSY INVASIVE DUCTAL CARCINOMA DUCTAL CARCINOMA IN SITU SEE TABLE 1 MULTIPLE LEVELS EXAMINED TABLE OF PATHOLOGICAL FINDINGS 1 INVASIVE CARCINOMA Tumor size <unk> X <unk> X 1 3cm Grade 2 **Lymphatic vessel invasion Present Blood vessel invasion Not identified** Margin of invasive carcinoma Invasive carcinoma extends to less than 0 2cm from the inferior margin of the specimen in one focus Location of ductal carcinoma in situ …	84%

MIT-MGH Validator



Modeling Disease Progression

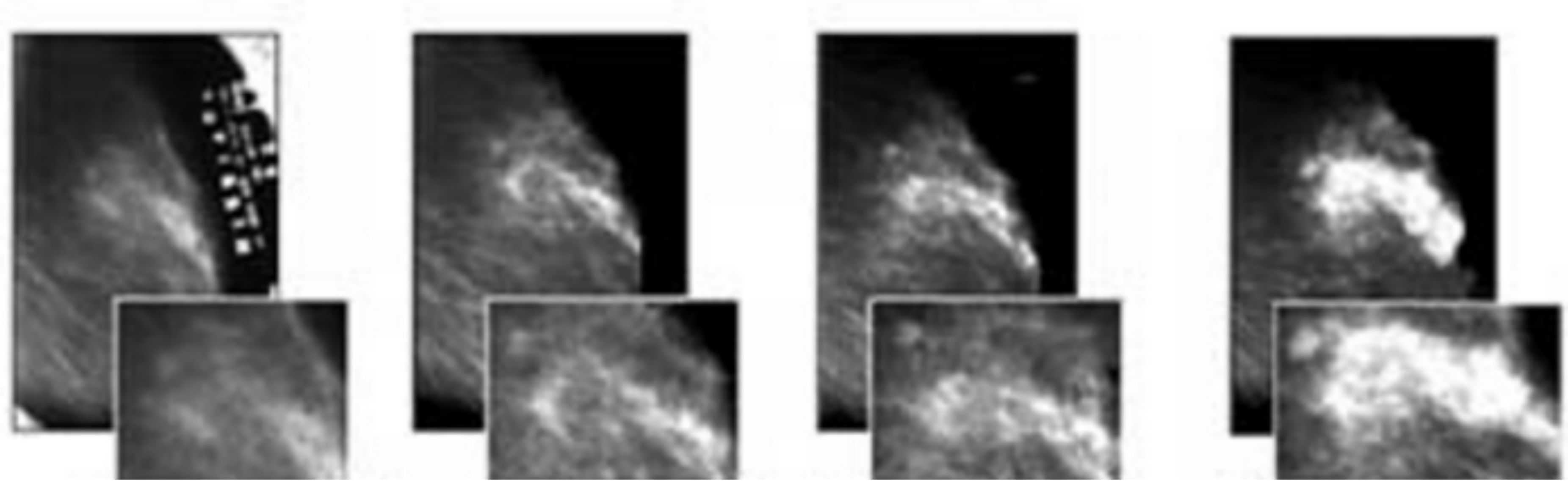




AST	8~ 40 U/L
ALT	5~ 40 U/L
LDH	50~400 U/L
ALP	80~280 U/L
γ-GTP	0~ 50 U/L
ZTT	4~ 12 KU
TTT	0~ 5 KU
Total bilirubin	0.4~0.9 mg/dℓ
Direct bilirubin	0~0.4 mg/dℓ
Indirect bilirubin	0.4~0.5 mg/dℓ
Total protein	6.0~9.0 g/dℓ

Predict Recurrence

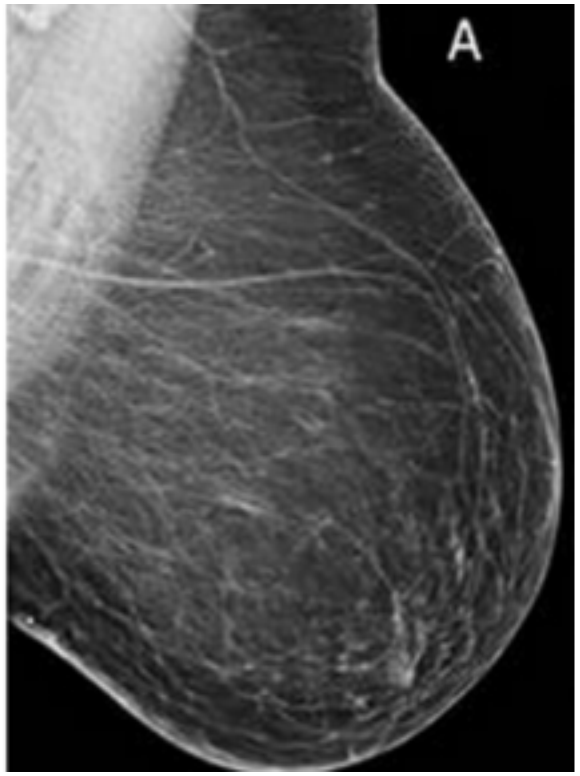
Early Cancer Detection

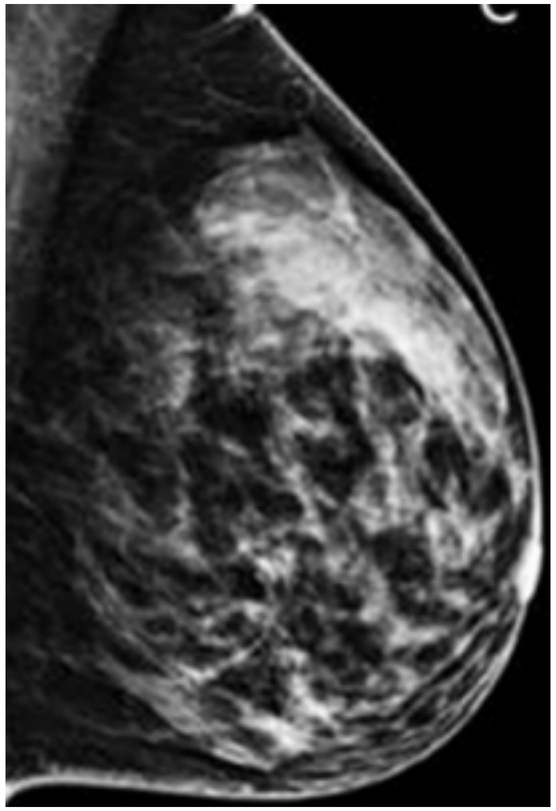


Hypothesis: changes in tissue are early cancer precursors

- support in clinical studies
- currently is not utilized in practice

Predicting Density

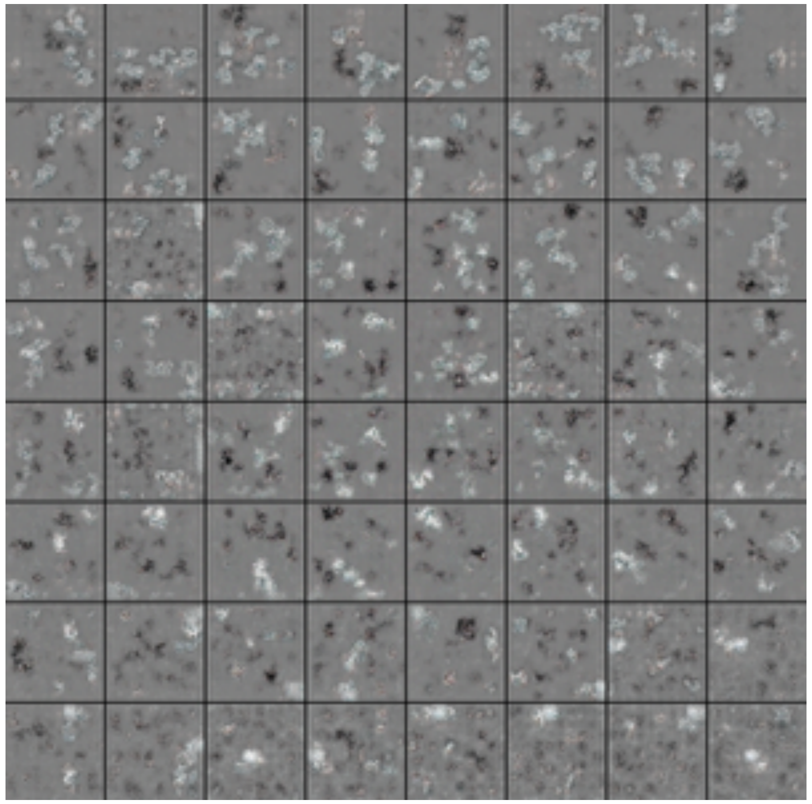




Density: ratio of fatty and fibrous tissue

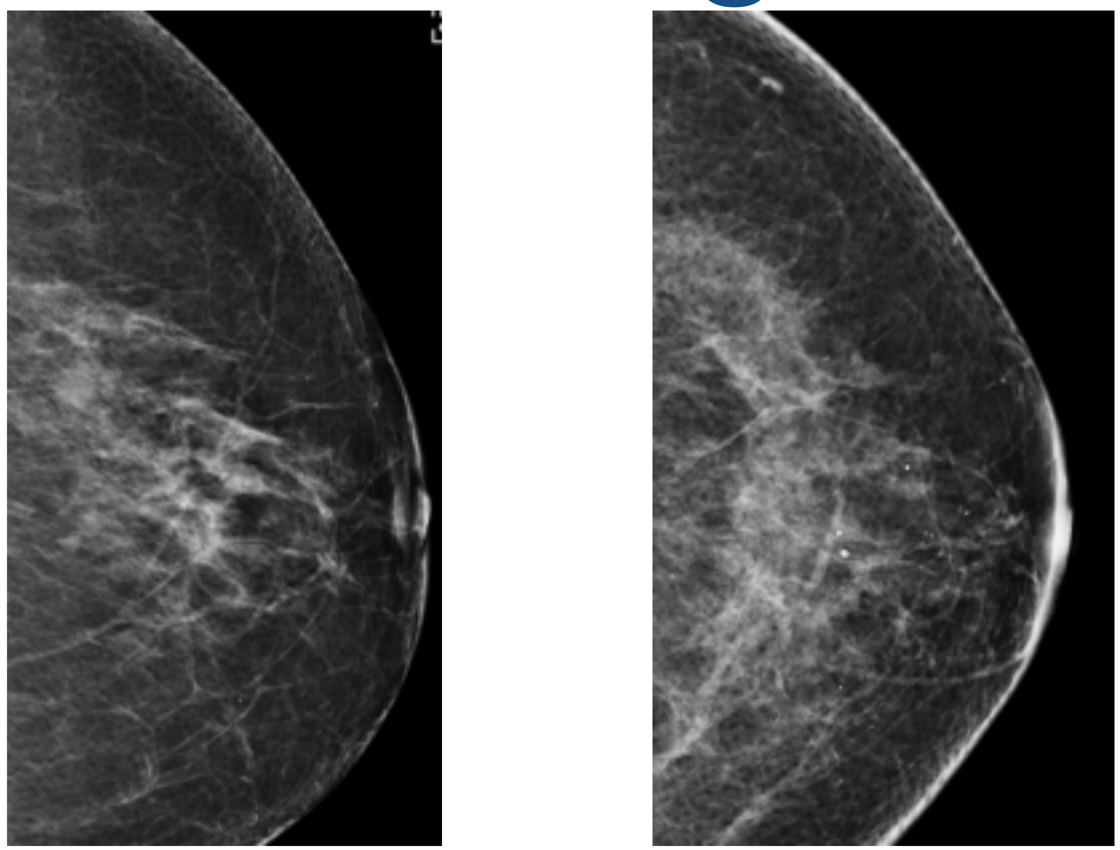
- increases lifetime breast cancer risk
- assessment is subjective

Predicting Density



- Training: 20K , Test: 5K
- Original 3K*3K images are downsampled to 256*256
- AlexNet Accuracy: 88%
- Human Agreement: 86%

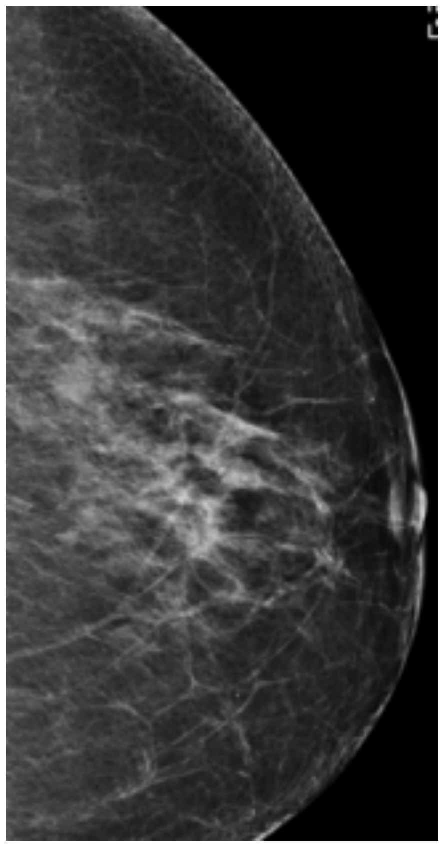
Predicting BI-RAD

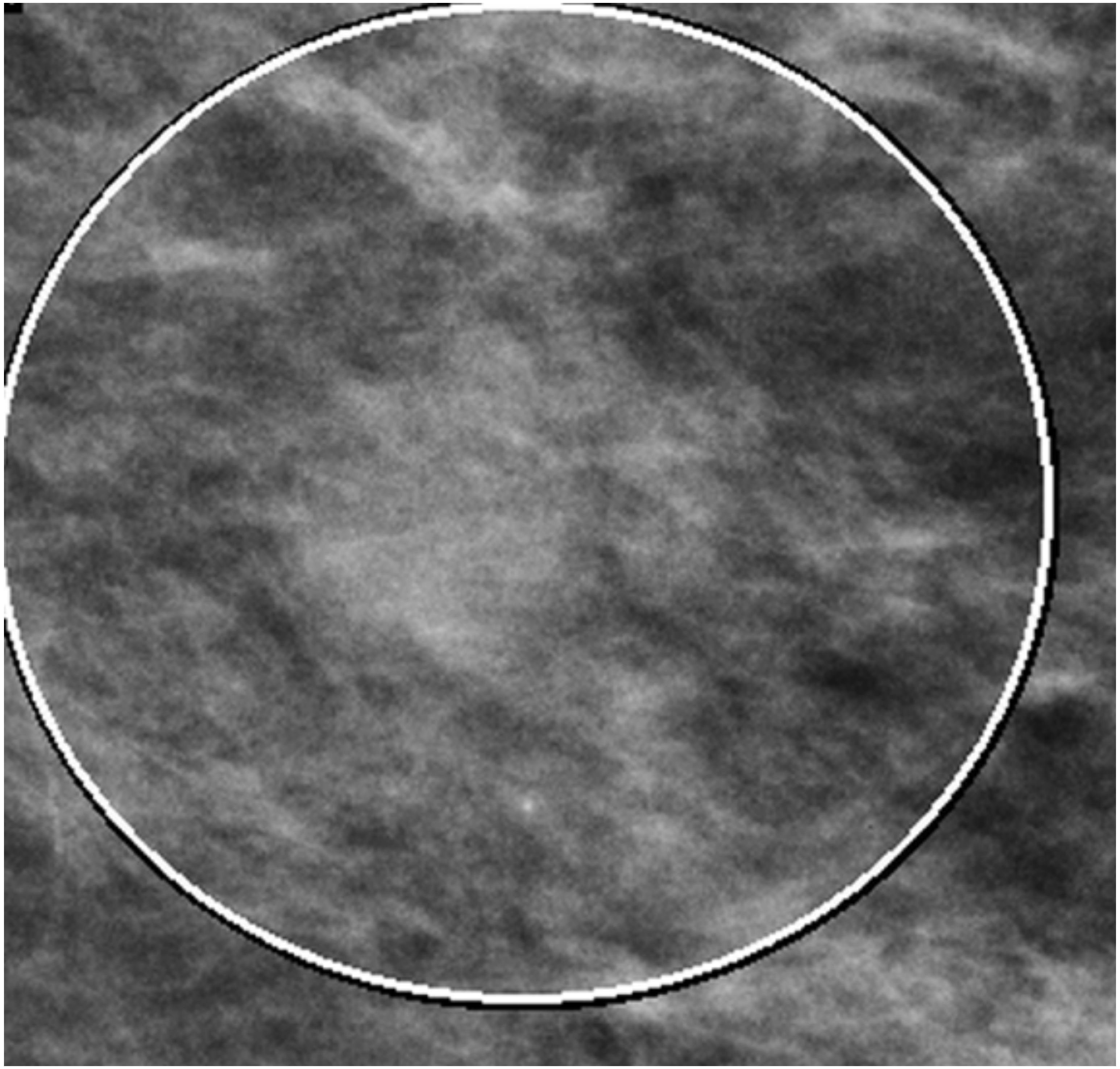


BI-RAD: quantifies the need in additional testing due to abnormality

- reasons: mass, asymmetry, distortion of tissue, …
- radiologists can assess BI-RAD with high accuracy

ZOOMED IN MAMMO





Predicting BI-RAD

- Train: 28K, Test: 1K
- Methods: AlexNet, Inception, Resnet, VGGNet, …
- F-measure: AUC 61%, F1 51%

Consistent with other reported results!

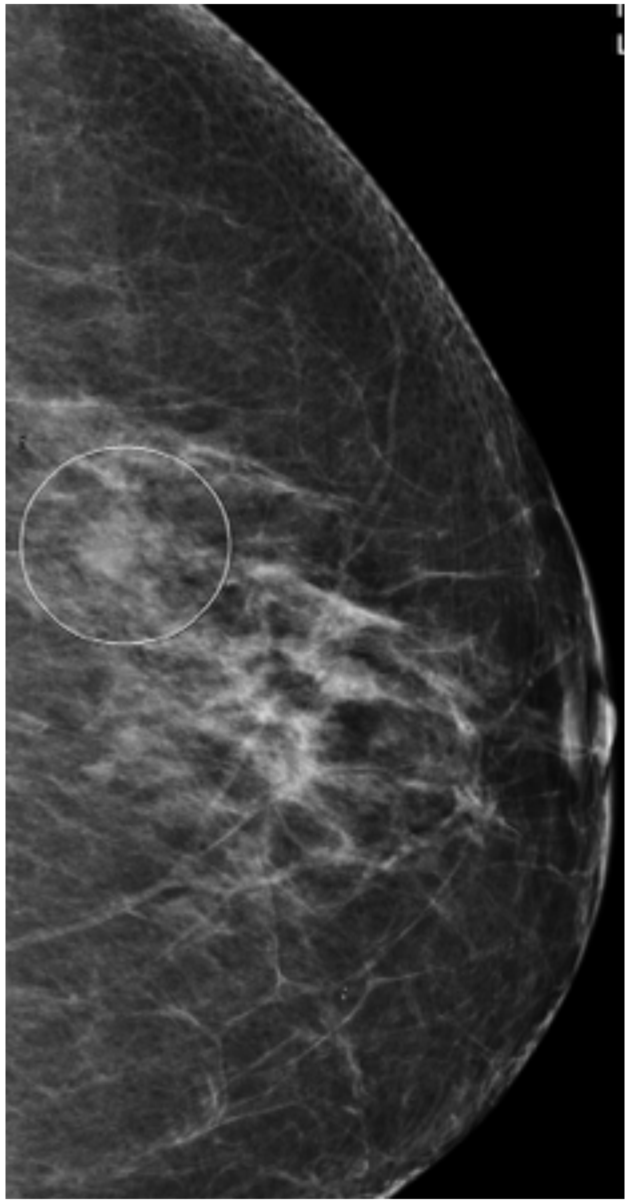
High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks

Krzysztof J. Geras[a], Stacey Wolfson[c], S. Gene Kim[c,d], Linda Moy[c,d], and Kyunghyun Cho[a,b,1]

[a]Center for Data Science, New York University; [b]Courant Institute of Mathematical Sciences, New York University; [c]Center for Biomedical Imaging, Radiology, NYU School of Medicine; [d]Perlmutter Cancer Center, NYU Langone Medical Center

Can Text Help to Focus Attention?

Finding 1: There is a focal asymmetry in the upper outer quadrant of the left breast. Finding 2: There are two asymmetries in the lateral aspect of the left breast. Finding 3: There are post lumpectomy changes in the right breast. IMPRESSION: Finding 1: Focal asymmetry in the upper outer quadrant of the left breast requires additional evaluation. Recommend additional mammographic views and ultrasound if warranted. Finding 2: Asymmetries in the lateral aspect of the left breast require additional evaluation.



Infinite thanks to my students and collaborators who helped to cure me