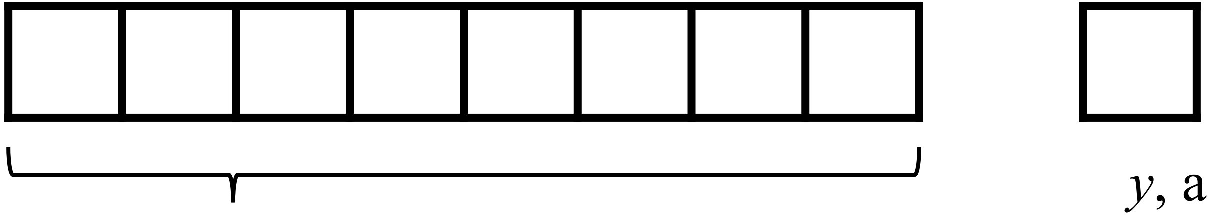
Introduction to Natural Language Processing in Healthcare

Matthew Engelhard

Lecture 1: what is a predictive model?



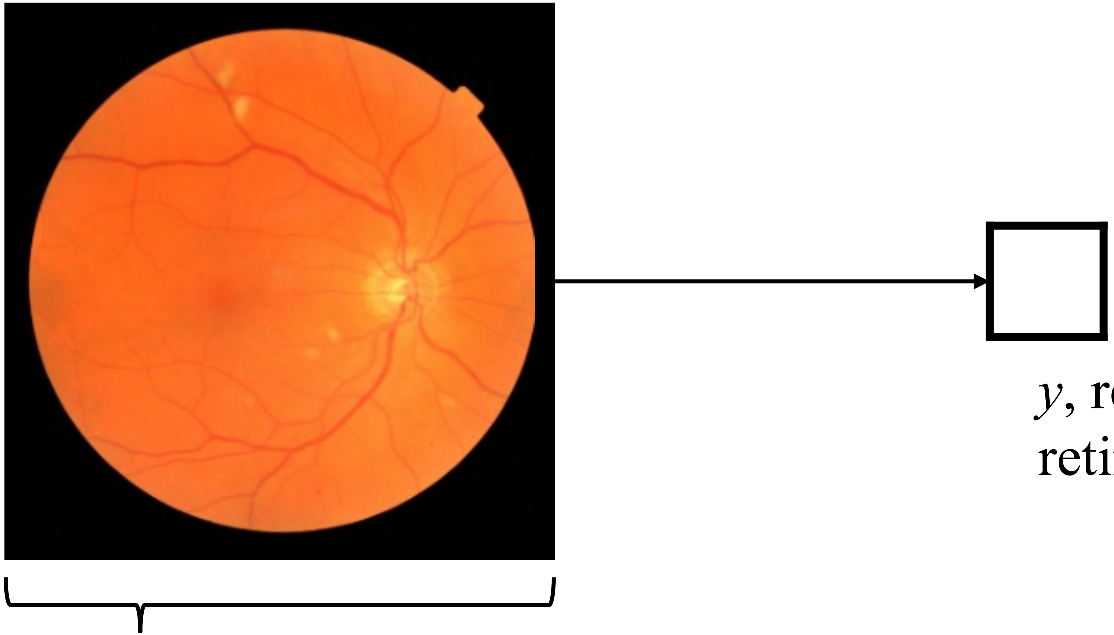
y, associated value or label

x, data/features for a subject or patient

End goal: predict *y* from *x*

-> Brief review

CNN: a predictive model for image data



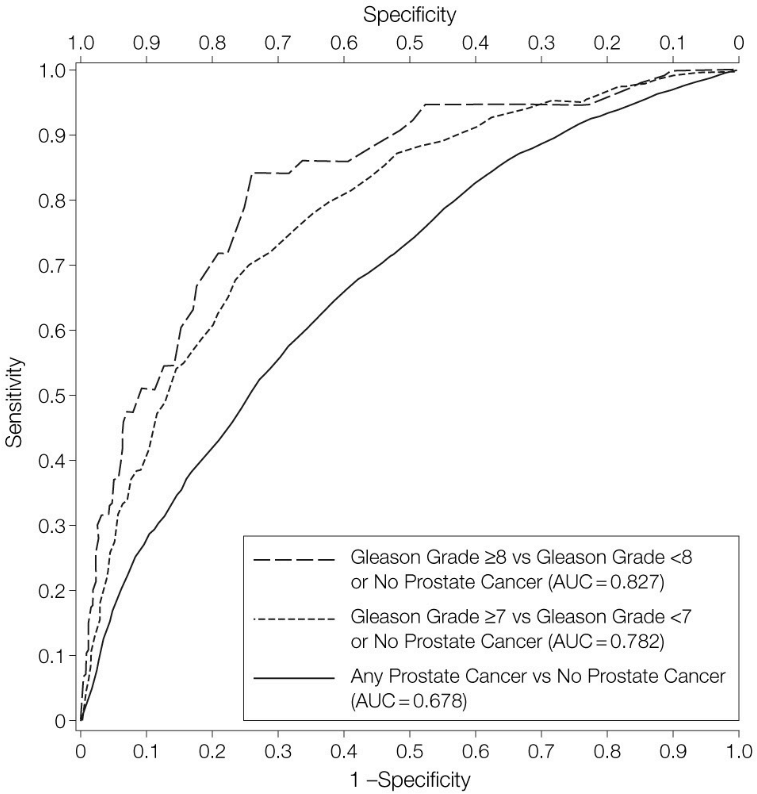
y, referrable diabetic retinopathy

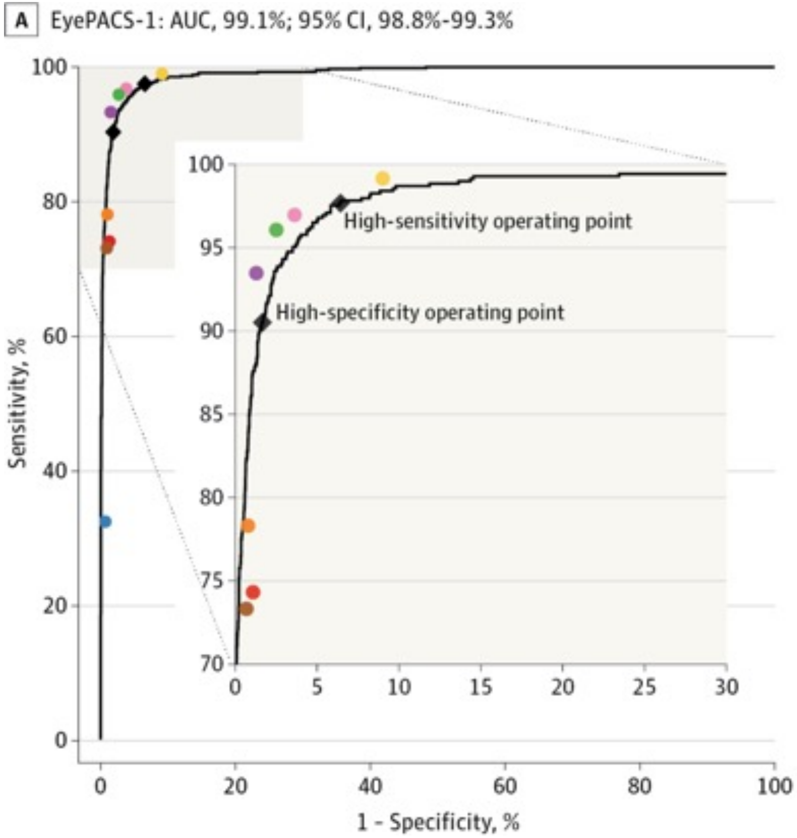
x, retinal image

End goal: predict *y* from *x*

-> Brief review

Evaluate performance just like any other diagnostic tool





-> Brief review

NLP: Predictive models for text data

- What can “natural language processing” (NLP) do?
  - Existing non-medical applications
  - Possible medical applications
- Binary and Multi-Class Classification can take place on:
  - The word level
  - The document or note level
  - The corpus or patient level
- Today: How can we convert text into something a predictive model can understand?

Text Generation

SYSTEM PROMPT (HUMAN-WRITTEN)

A train carriage containing controlled nuclear materials was stolen in Cincinnati today. Its whereabouts are unknown.

"Better Language Models and Their Implications"
2/14/19
OPENAI

MODEL COMPLETION (MACHINE-WRITTEN, FIRST TRY)

The incident occurred on the downtown train line, which runs from Covington and Ashland stations.

In an email to Ohio news outlets, the U.S. Department of Energy said it is working with the Federal Railroad Administration to find the thief.

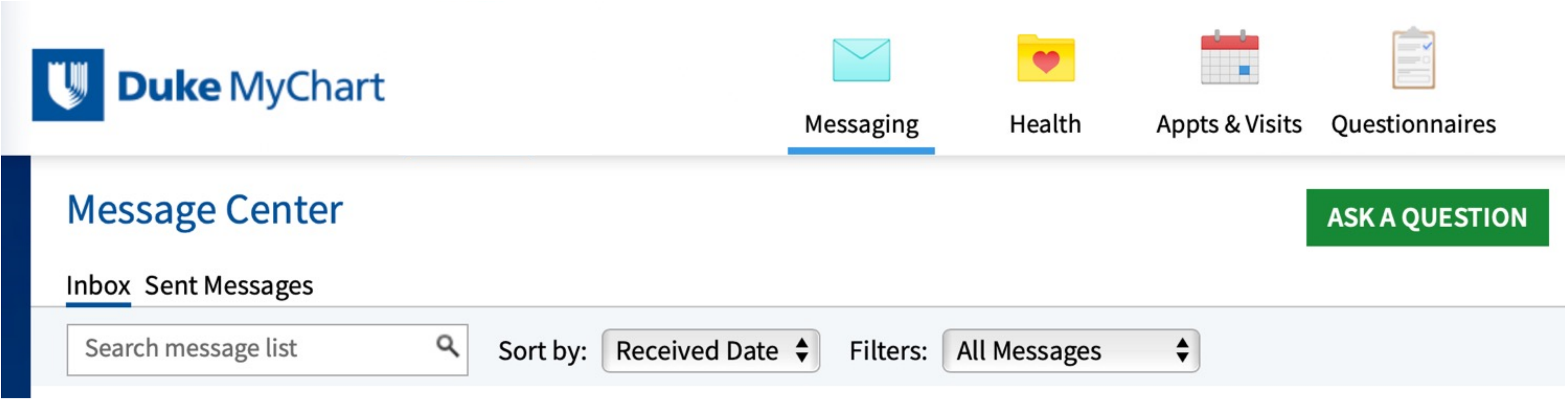
"The theft of this nuclear material will have significant negative consequences on public and environmental health, our workforce and the economy of our nation," said Tom Hicks, the U.S. Energy Secretary, in a statement. "Our top priority is to secure the theft and ensure it doesn't happen again."

The stolen material was taken from the University of Cincinnati's Research Triangle Park nuclear research site, according to a news release from Department officials.
The Nuclear Regulatory Commission did not immediately release any information.

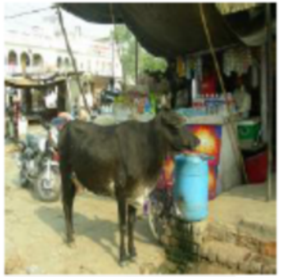
According to the release, the U.S. Department of Energy's Office of Nuclear Material Safety and Security is leading that team's investigation.

"The safety of people, the environment and the nation's nuclear stockpile is our highest priority," Hicks said. "We will get to the bottom of this and make no excuses.

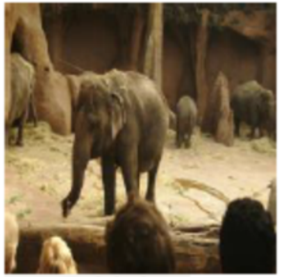
Suggested Email Responses



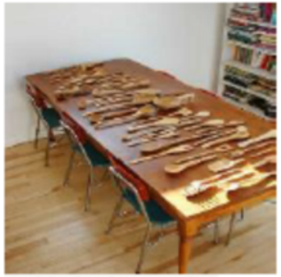
Automatic Image Captioning



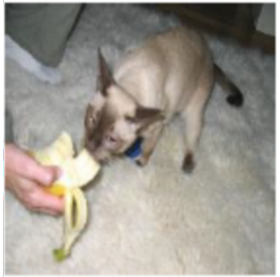
a cow is standing in front of a store



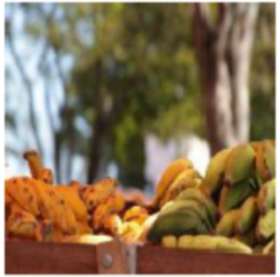
a group of elephants standing next to each other



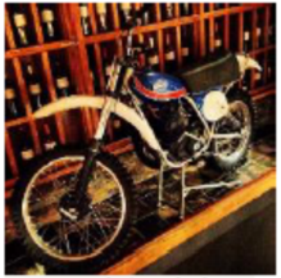
a table that has wooden spoons on it



a cat is eating some kind of food

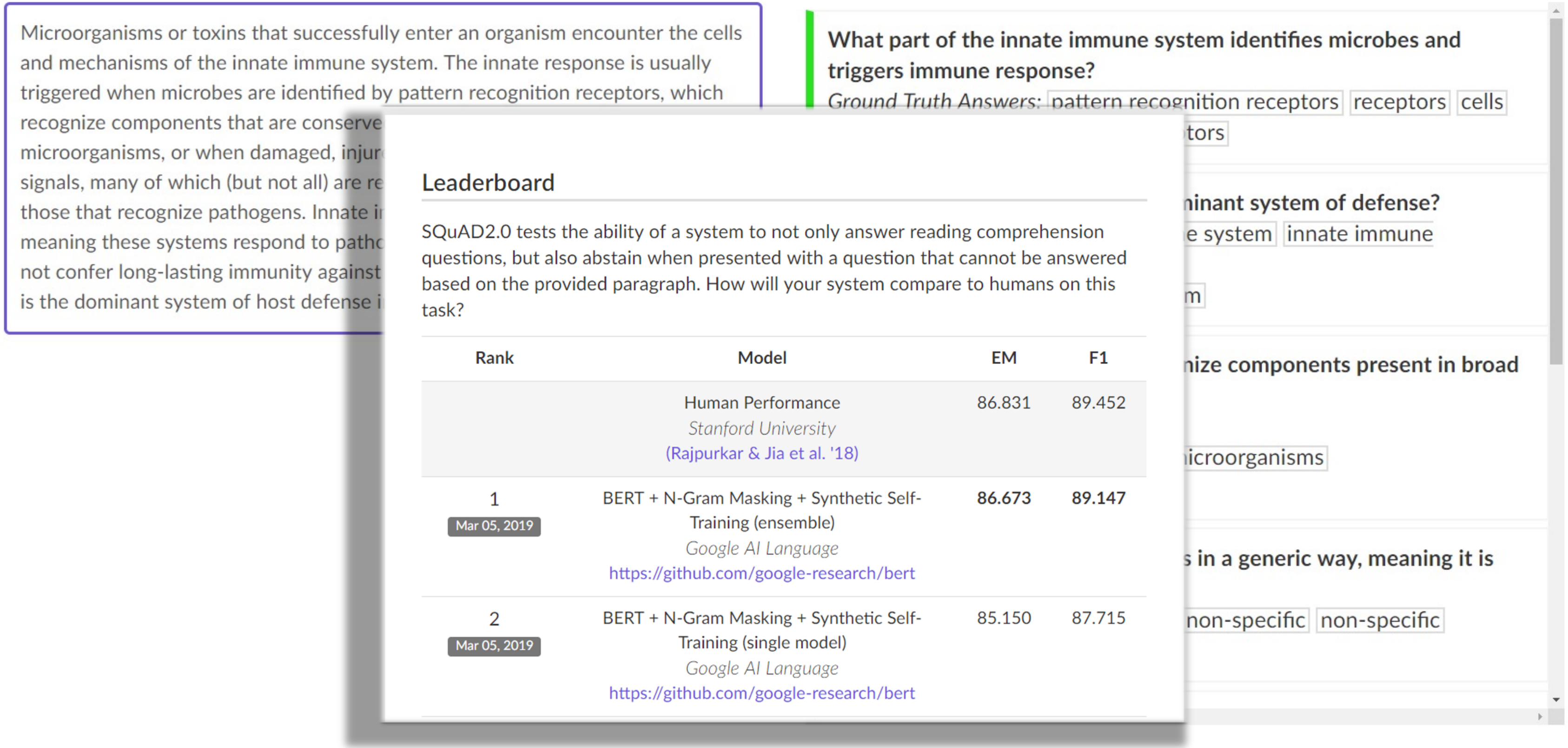


a bunch of bananas are sitting on a table



a motorcycle is parked next to a window

Question Answering



Microorganisms or toxins that successfully enter an organism encounter the cells and mechanisms of the innate immune system. The innate response is usually triggered when microbes are identified by pattern recognition receptors, which
recognize components that are conserve
microorganisms, or when damaged, injur
signals, many of which (but not all) are re
those that recognize pathogens. Innate i
meaning these systems respond to patho
not confer long-lasting immunity against
is the dominant system of host defense i

What part of the innate immune system identifies microbes and triggers immune response?

Ground Truth Answers: pattern recognition receptors | receptors | cells

tors

ninant system of defense?

e system | innate immune

m

nize components present in broad

icroorganisms

s in a generic way, meaning it is

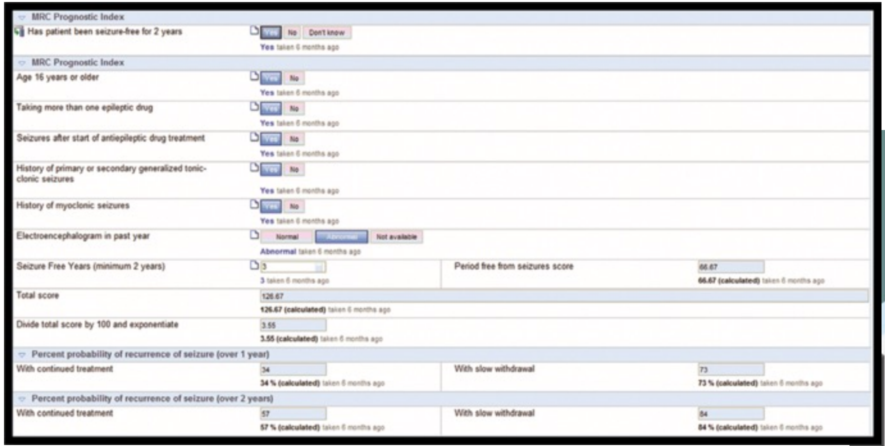
non-specific | non-specific

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph. How will your system compare to humans on this task?

Rank	Model	EM	F1
	Human Performance *Stanford University* (Rajpurkar & Jia et al. '18)	86.831	89.452
1 Mar 05, 2019	BERT + N-Gram Masking + Synthetic Self-Training (ensemble) *Google AI Language* https://github.com/google-research/bert	**86.673**	**89.147**
2 Mar 05, 2019	BERT + N-Gram Masking + Synthetic Self-Training (single model) *Google AI Language* https://github.com/google-research/bert	85.150	87.715

Populating Standardized Forms



MRC Prognostic Index

Has patient been seizure-free for 2 years	[✓ Yes] No Don't know **Yes** taken 6 months ago		

MRC Prognostic Index

Age 16 years or older	[✓ Yes] No **Yes** taken 6 months ago		
Taking more than one epileptic drug	[✓ Yes] No **Yes** taken 6 months ago		
Seizures after start of antiepileptic drug treatment	[✓ Yes] No **Yes** taken 6 months ago		
History of primary or secondary generalized tonic-clonic seizures	[✓ Yes] No **Yes** taken 6 months ago		
History of myoclonic seizures	[✓ Yes] No **Yes** taken 6 months ago		
Electroencephalogram in past year	Normal [✓ Abnormal] Not available **Abnormal** taken 6 months ago		
Seizure Free Years (minimum 2 years)	3 3 taken 6 months ago	Period free from seizures score	66.67 **66.67 (calculated)** taken 6 months ago
Total score	126.67 **126.67 (calculated)** taken 6 months ago		
Divide total score by 100 and exponentiate	3.55 **3.55 (calculated)** taken 6 months ago		

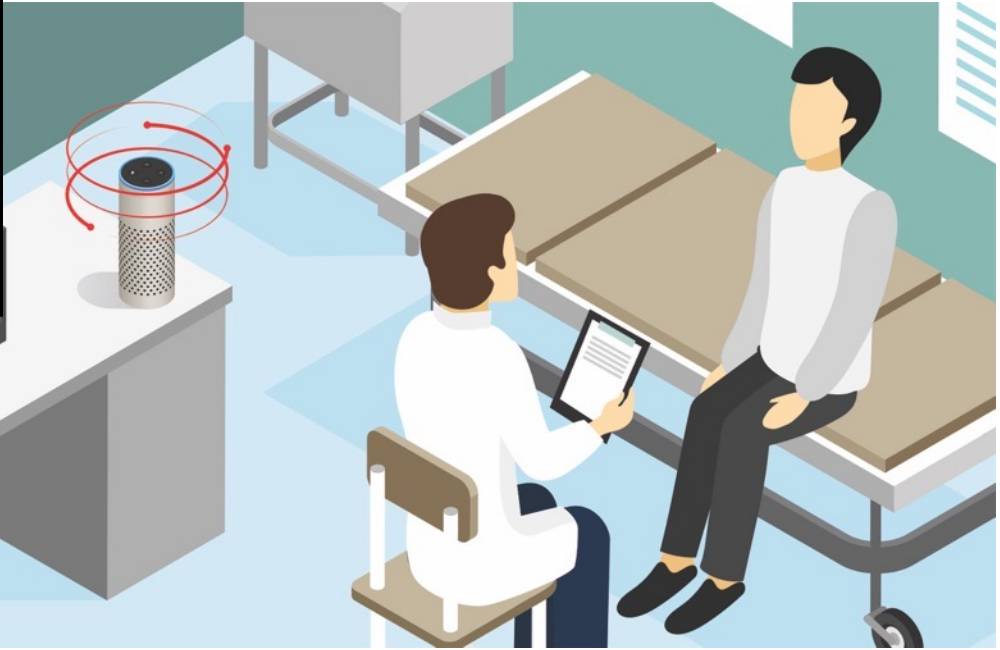
Percent probability of recurrence of seizure (over 1 year)

With continued treatment	34 **34 % (calculated)** taken 6 months ago	With slow withdrawal	73 **73 % (calculated)** taken 6 months ago

Percent probability of recurrence of seizure (over 2 years)

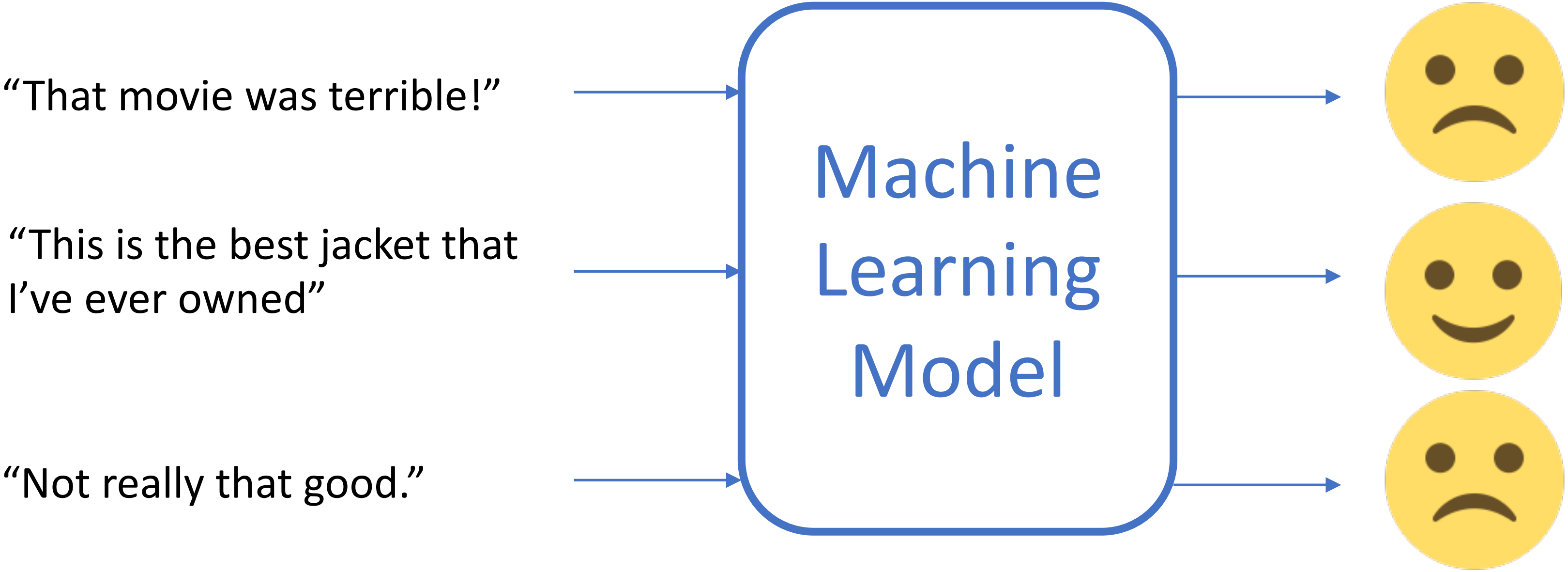
With continued treatment	57 **57 % (calculated)** taken 6 months ago	With slow withdrawal	84 **84 % (calculated)** taken 6 months ago

Narayanan et al,
Epilepsia (2017)



Our Focus: Classification.

For example, sentiment analysis



“That movie was terrible!”

“This is the best jacket that I’ve ever owned”

“Not really that good.”

Machine Learning Model

Binary Classification of Documents

- Food or movie reviews (positive vs negative)
- Clinical notes with a specific finding

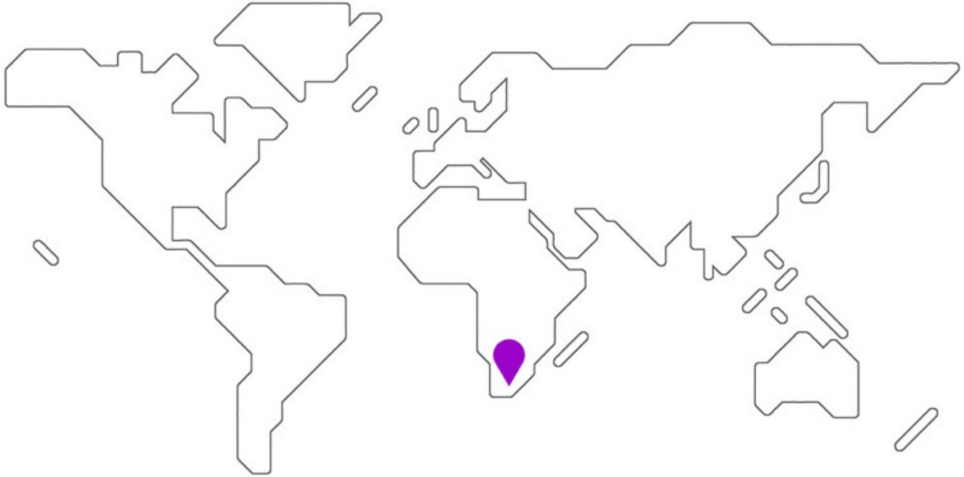
Multi-Class Classification of Documents

- Food or movie reviews (scored)
- Clinical notes with several types of content of interest
- Findings in radiology or surgery reports

Case Study: SMS Triage for Global Maternal Health

Maternal Health HelpDesk:

2 million women connected to NDoH staff via SMS



https://www.praekelt.org

Binary Classification: Urgent Message? (Yes/No)

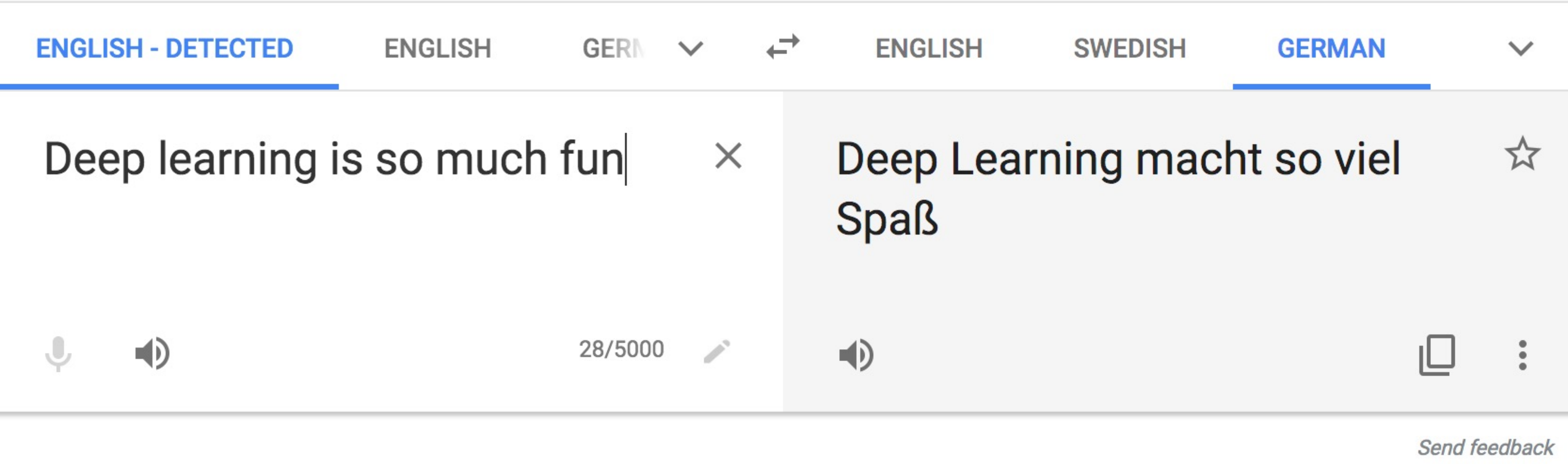
Word-level classification

- De-identification of patient notes
- Identification of specific medical terms and concepts
- Move information from free-text to structured fields

There are a few ambiguous cases:

- Question answering kind of belongs here…

Text Translation



Deep learning is so much fun

Deep Learning macht so viel Spaß

translate.google.com

Corpus-level (or **patient**-level) classification

- Diagnosis prediction, readmission prediction, etc

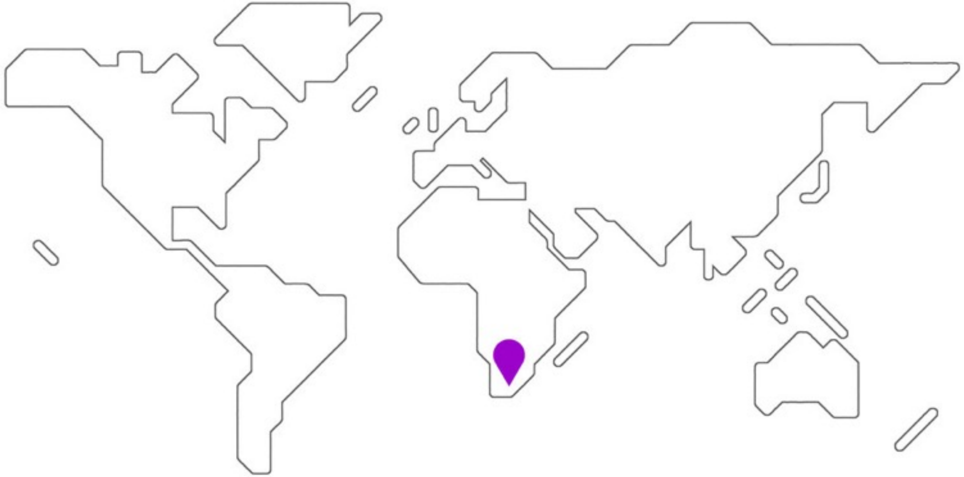
On this level, we start running into practical challenges:

- Low signal to noise
- Too much data to fit into memory
- May need to train using a subset of all notes / documents
  - e.g. select all discharge summaries or all notes from a particular specialty

Case Study: SMS Triage for Global Maternal Health

Maternal Health HelpDesk:

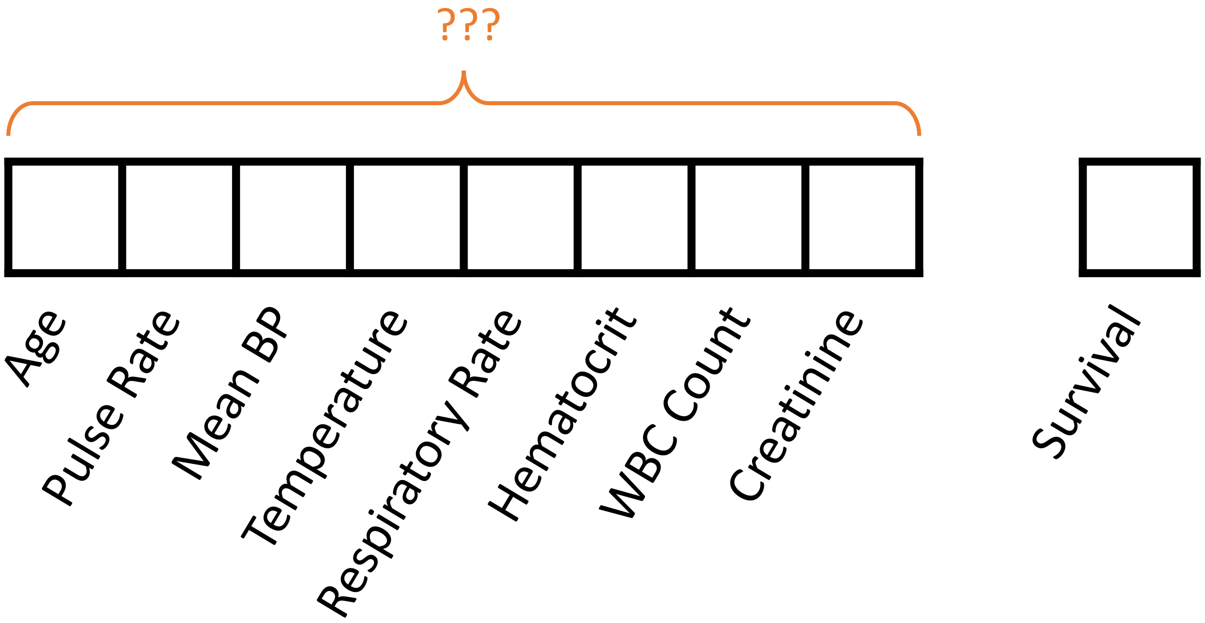
2 million women connected to NDoH staff via SMS



https://www.praekelt.org

Binary Classification: Urgent Message? (Yes/No)

A Simple Predictive Model: ICU Mortality



End goal: predict odds of hospital mortality

We need numbers, not words

- **Can we convert our text to a vector or sequence of numbers?**
- If yes, we can use logistic regression (or any other predictive model)!
- <u>Next Lecture</u>: How to convert text to numeric features