AAMC CENTER FOR HEALTH JUSTICE

MATERNAL HEALTH EQUITY WORKSHOP:
FROM STORY TO DATA TO ACTION
MAY 18, 2023

Maternal Health Equity Workshop: From Story to Data to Action

Association of
American Medical Colleges

Bianca's Story: Turning Pain into Action

Bianca Dickerson-Williams, JD
Founder and Executive Director, Fighting Our Injustices for Women of Color

Fighting our Injustices

AAMC CENTER FOR HEALTH JUSTICE

Background

This story needs to be heard to save the lives of Black women, Black children, and Black men.

Black Maternal Health Crisis: A Genocide of the African American Race/ Systemic Racism.

Background of my story. (Misconception about who and the stereotypes)

Law Enforcement and Lawyer.

Factual Details of the case and what happened in the delivery room.

Statistics: Department of Health and Center for Disease Control

The United States has the highest maternal mortality rate amongst developed countries.

Black infants are twice as likely to die in their first year of life than White infants.

In California, Black women are six times more likely to die due to complications within the first year of pregnancy than White women.

In the United States, Black women are three times more likely to die from pregnancy complications.

Black women have a 70% increased risk for severe maternal morbidity.

Indigenous women have the next highest morbidity and maternal health issues.

Police brutality and healthcare brutality

Why do they conceal (insurance, credibility, stain on their career, medical license suspended or revoked).

Doctors are more important than us and they have been put on a pedestal by society.

Police body cameras versus no other checks and balances for physicians. Their story is the story regardless if it is true or not.

Abuse of power: “whatever it is they say happened, that is what happened”

Corruption

Concealment of issues due to a lack of documentation, altering the truth, or omitting the truth.

Altered medical records, deletions, and omissions/failure to truthfully document in police reports and medical records

Doctors cover for other doctors just like police cover for other police. It is a systemic abuse of power.

Call to action

Momnibus Legislation

Women of Color in Policy Coalition

U.S. Congress (Making laws to stop the problem)

Legal Matters

The White House

Documentary “Bianca’s Story: Black Moms At Risk”

Youtube and Fightingourinjustices.com

AAMC CENTER FOR HEALTH JUSTICE

The AAMC Center for Health Justice Welcomes You to the Maternal Health Equity Workshop: From Story to Data to Action

Philip M. Alberti, PhD
Senior Director, Health Equity Research & Policy
Founding Director, AAMC Center for Health Justice

Carla S. Alvarado, PhD, MPH
Director of Research, AAMC Center for Health Justice

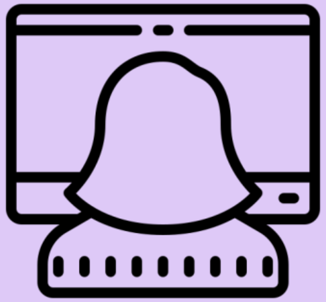
Agenda

10:20 a.m.	From Story to Data: Understanding Natural Language Processing
10:50 a.m.	From Data to Research: NLP & Maternal Health Use Cases
11:45 a.m.	Break
12:00 p.m.	Beyond Buzzwords: Reimagining the Default Settings of Technology & Society
12:35 p.m.	Exercise: Principles of Trustworthy NLP use in Maternal Health
2:00 p.m.	From Data to Action: What Public Health, Hospitals and Health Systems Can Do
2:50 p.m.	Drawing Change Workshop Summary

***All times are ET**

Introduction to NLP

from story to data to action



Maria Antoniak, PhD

Allen Institute for AI

AAMC Maternal Health Equity Workshop, May 18th 2023

Welcome!

During this session, we're going to cover the basics of NLP: methods, models, and challenges.

NLP comes with both big risks and exciting potential.

You can better understand how to weigh these risks and benefits if you know how NLP works "under the hood."

I'm excited to learn with you today!

Our goals

1. Overview of different methods and workflows

2. Breadth not depth; the other sessions will go deeper

3. Understand the assumptions underlying these models

4. Awareness of how text turns into numbers used by NLP systems

About me

I’ve been working in NLP for a decade both in academia and industry.

Right now, I’m a postdoc at the Allen Institute for AI. We’re a research nonprofit, and our goal is to study “AI for the common good.”

I have a lot of worries about AI but also a lot of hope. I want to open up AI so that as many different people and perspectives are included as possible.

Teaching you the basics of NLP is part of that goal! **We need your expertise.**

Overview

What is natural language processing?

NLP, or **computational linguistics**, uses computational methods to study human language.

This can include **analyzing** human language and **generating** human language.

Methods can include statistics, machine learning, linguistics, and programming.

Examples:

- Google Search
- Alexa
- Email spam filters
- Autocomplete
- ChatGPT

Why are people excited about NLP?

1. Human language is fascinating!

2. We can do so many different things using NLP!
    - question answering
    - measuring biases
    - extracting information
    - … and more!

    To see lots of exciting examples, make sure you attend all of the talks today!

3. Recent advances makes more things possible than most of us expected!

Why are people wary or critical of NLP?

Models can encode **biases** that are difficult to measure or correct.

Many systems rely on **poorly documented training data**.

Models are very large and training them takes up a lot of **resources**.

The processes leading to results can be hard to **interpret**.

For more on this topic, make sure you attend Dr. Ruha Benjamin's talk later today!

From story to data

Computers don't understand words, but they do understand numbers!

How can we turn **words** into **numbers?**

language
technology
baby
pregnancy
maternal health
computer

37
0
41
102
18
24
94
346
55

“You shall know a word by the company it keeps.”

The **distributional hypothesis**, popularized by Firth (1957).

“You shall know a word by the company it keeps.”

In other words: *Which words often appear together?*

We can learn about the **meaning** of a word by studying its **usage**.

In NLP, we often use a word’s usage patterns as a **proxy** for its **semantic relationships** to other words.

Definition Alert

Vector

A list of numbers

8	32	0	0	3	99

Matrix

A list of lists

8	32	0	0	3	99
74	0	5	45	0	0
3	26	0	5	0	92
0	0	16	0	64	23

Turning words into numbers

Use the distributional hypothesis!

Create a **matrix** where each cell contains the number of times the words occur in the same sentence.

	computer	technology	maternal	baby
computer	10	5	2	0
technology	5	20	1	1
maternal	2	1	11	6
baby	0	1	6	13

Turning words into numbers

Each row of the matrix is a **vector** that represents the word.

`baby = [10, 5, 2, 0]`

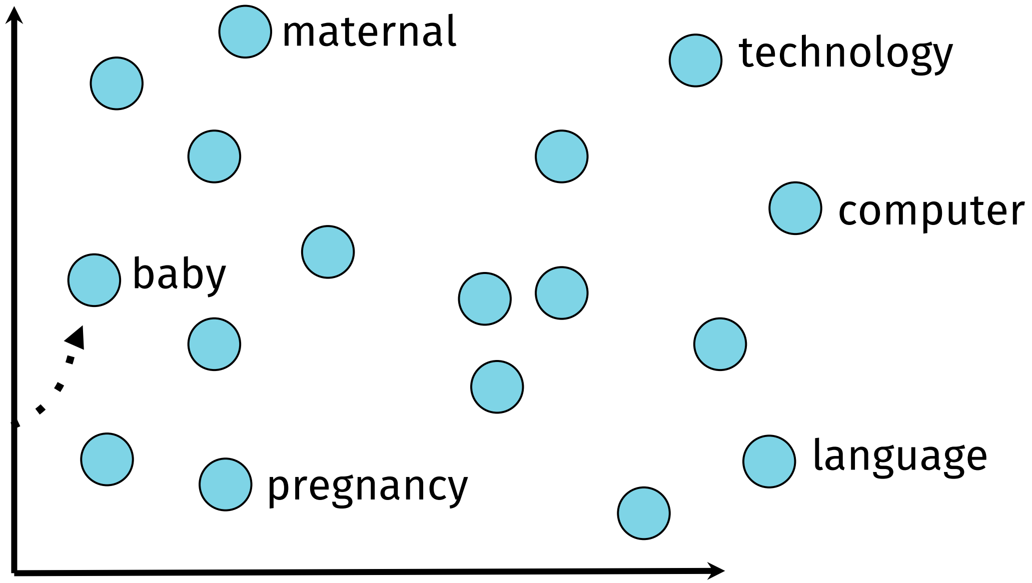
	computer	technology	maternal	baby
computer	10	5	2	0
technology	5	20	1	1
maternal	2	1	11	6
baby	0	1	6	13

What can we do with those numbers?

We can plot the words in a graph.

And then we can measure distances and relationships between words!

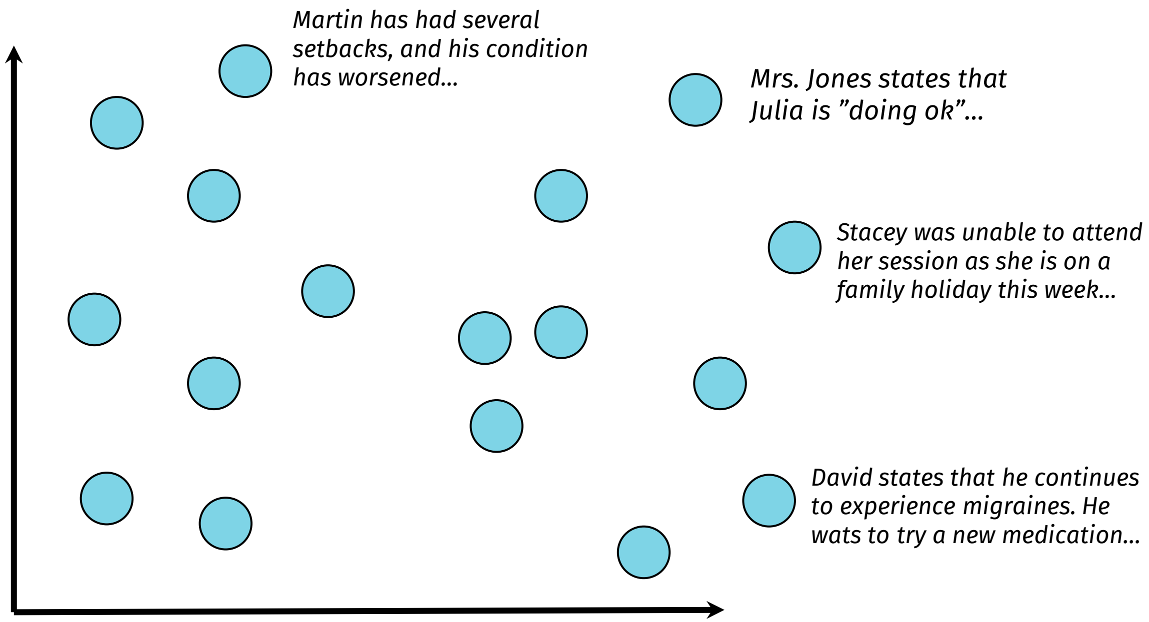
```
baby = [10, 5, 2, 0]
```



What can we do with those numbers?

For whole texts, we can average all the word vectors to get a single vector for the text.

Now we can measure distances and relationships between the texts.



Word Vectors (or Word Embeddings)

There are many different ways to create these word vectors.

Latent semantic analysis (LSA) is a simple method from 2004.

In 2013, **word2vec** was introduced, which used a neural method to discover the set of vectors.

Where do we get all those word counts?

These models take advantage of **big datasets** and **pretraining** on a large dataset and applying that model to a smaller task.

Datasets can include:

- Wikipedia
- Reddit
- books
- EHR
- scientific publications

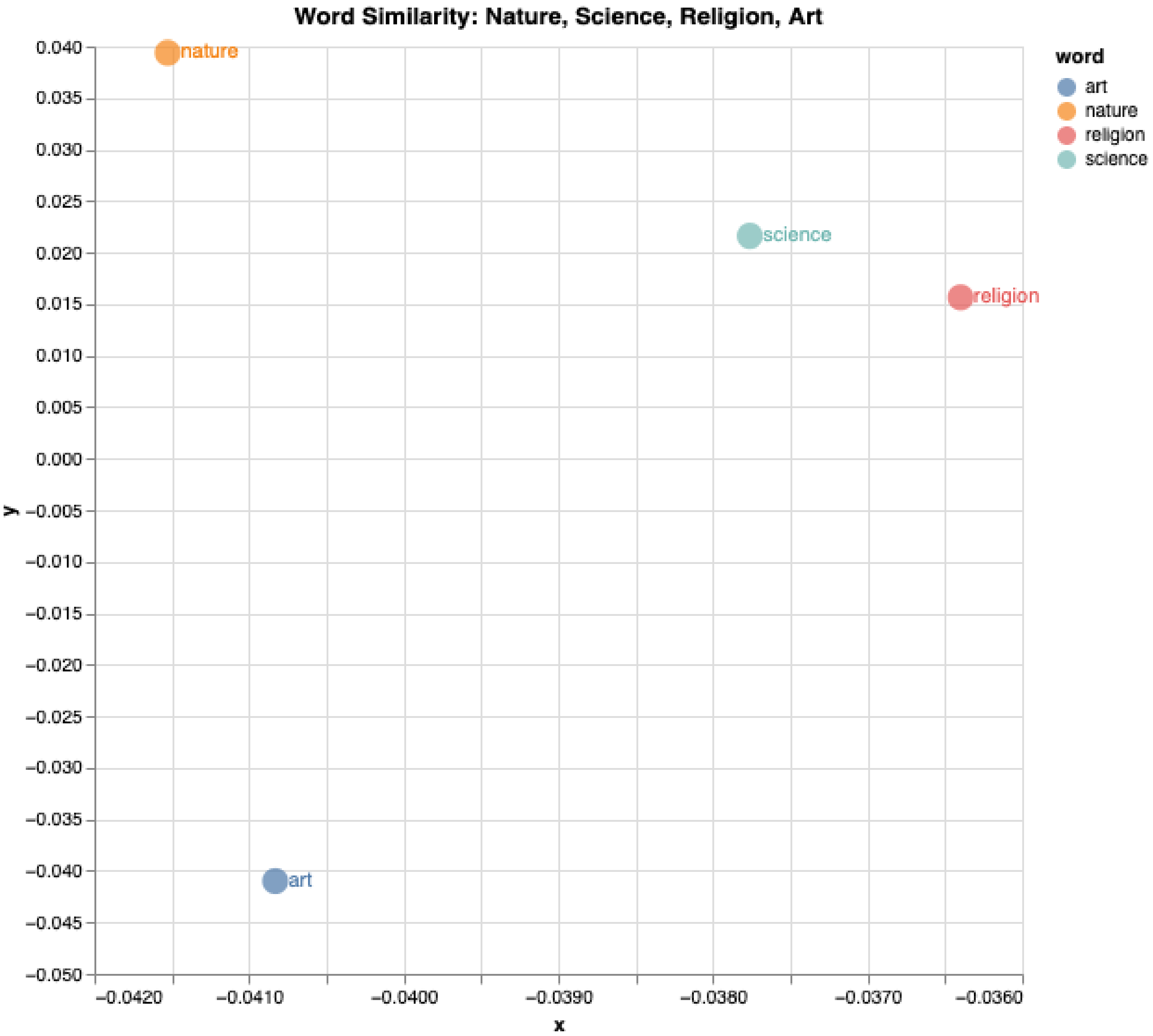
Contextualized vectors

Word2vec and similar methods produce **static** vectors, where each word is represented by a single vector.

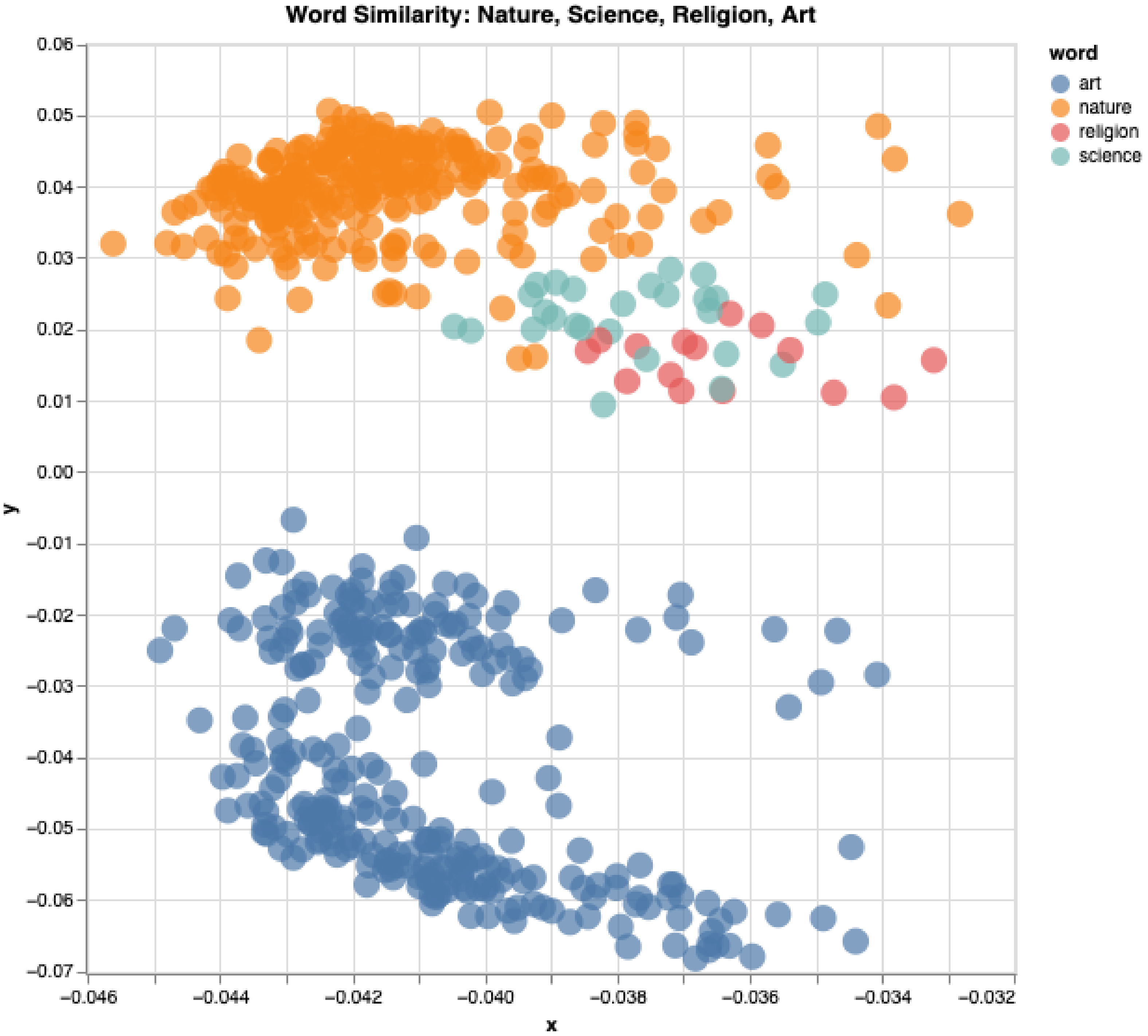
Newer methods create a new vector for each time a word is used in a dataset. These are **contextualized** vectors.

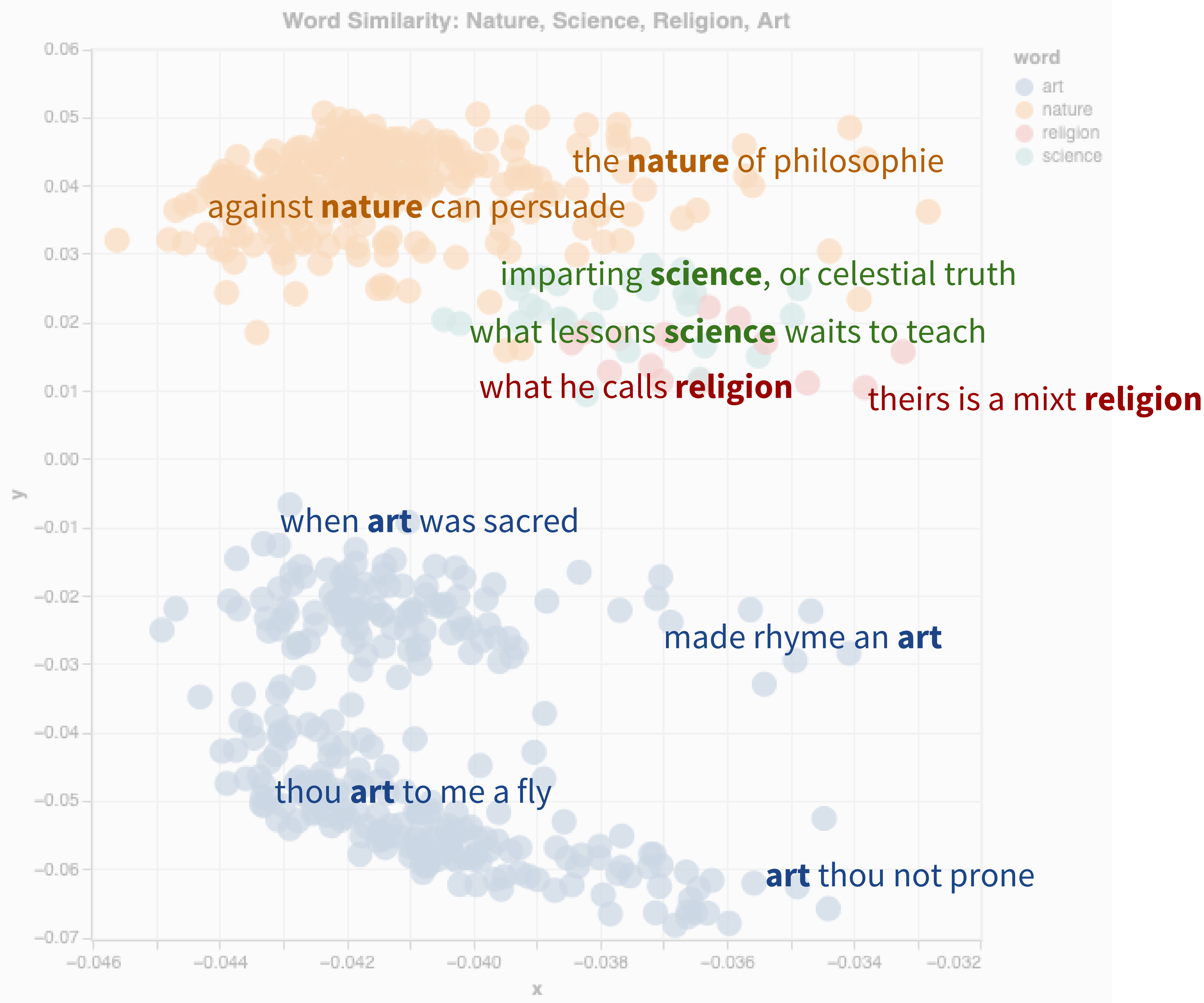
These are some of the methods enabling new tools like ChatGPT.

Static Vectors: one point per vocabulary item



Contextualized Vectors: one point per instance of a word in context!





Word Similarity: Nature, Science, Religion, Art
word
art
nature
religion
science
the nature of philosophie
against nature can persuade
imparting science, or celestial truth
what lessons science waits to teach
what he calls religion
theirs is a mixt religion
when art was sacred
made rhyme an art
thou art to me a fly
art thou not prone
0.06
0.05
0.04
0.03
0.02
0.01
0.00
−0.01
−0.02
−0.03
−0.04
−0.05
−0.06
−0.07
y
−0.046
−0.044
−0.042
−0.040
−0.038
−0.036
−0.034
−0.032
x

Tokenization: Example

“Extempore Effusion upon the Death of James Hogg”

By William Wordsworth

When first, descending from the moorlands,

I saw the Stream of Yarrow glide

Along a bare and open valley,

The Ettrick Shepherd was my guide.

When last along its ***banks*** *I wandered,*

Through groves that had begun to shed

BERT Tokenization: Example

[CLS] when first , descending from the moor **##**lands , i saw the stream of ya **##**rrow glide along a bare and open valley , the et ##trick shepherd was my guide . when last along its banks i wandered , through groves that had begun to shed their golden leaves upon the pathways , my steps the border - min ##strel led . the mighty min **##**strel breathe **##**s no longer , \' mid mo ##uld ##ering ruins low he lies ; and death upon the bra **##**es of ya **##**rrow , has closed the shepherd - poet \' s eyes : nor has the rolling year twice measured , from sign to sign , its ste ##df ##ast course , since every mortal power of cole ##ridge was frozen at its marvel **##**lous source ; the rap **##**t one , of the god **##**like forehead , the heaven - eyed creature sleeps in earth : and lamb , the fr ##olic and the gentle , has vanished from his lonely hearth . like clouds that rake the mountain - summit **##**s , or waves that own no curb ##ing hand , how fast has brother followed brother , from sunshine to the sun **##**less land ! yet i , whose lids from infant sl ##umber were earlier raised , remain to hear a tim ##id voice , that asks in whispers , " who next will drop and disappear ? " our ha ##ught ##y life is crowned with darkness , like london with its own black wreath , on which with thee , o crab **##**be ! forth - looking , i gazed from hampstead \' s bree ##zy heath . as if but yesterday departed , thou too art gone before ; but why , o \' er ripe fruit , season ##ably gathered , should frail survivors he ##ave a sigh ? mo ##urn rather for that holy spirit , sweet as the spring , as ocean deep ; for her who , er ##e her summer faded , has sunk into a breathless sleep . no more of old romantic sorrow ##s , for slaughtered youth or love - lo **##**rn maid ! with sharpe **##**r grief is ya **##r**row sm **##**itte **##**n , and et **##**trick mo **##**urn **##**s with her their poet dead . **[SEP] [PAD]**

From data to model

Supervised vs Unsupervised Machine Learning

Supervised: texts have labels, model learns patterns from **texts + labels**

DATA	LABEL
The final score between the Patriots and the Seahawks...	Sports
The politician’s fiscal policies are untenable...	Op-Ed
Brian Jones was injured during the final inning...	Sports

Unsupervised: texts do **not** have labels, model learns patterns from **text only**

DATA	LABEL
The final score between the Patriots and the Seahawks...	?
The politican’s fiscal policies are untenable...	?
Brian Jones was injured during the final inning...	?

Supervised Methods: Examples

Let’s take a look at some supervised machine learning models!

These models assume that you have labels for your data, and you want to learn how to apply those labels to new data.

Supervised Methods: Examples

We have a dataset of texts and labels.

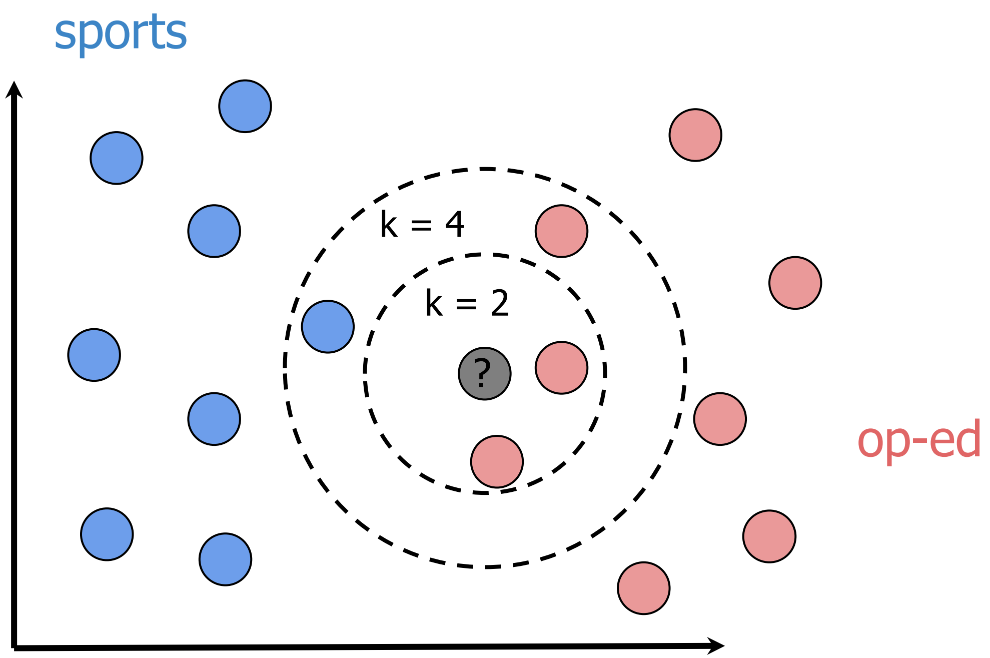
We divide that dataset into

- a **training** set (used to teach our model)
- a **test** set (used to test whether our model works on new data)

Learn from your neighbors

Intuition: assign the most common label amongst an input's k nearest neighbors

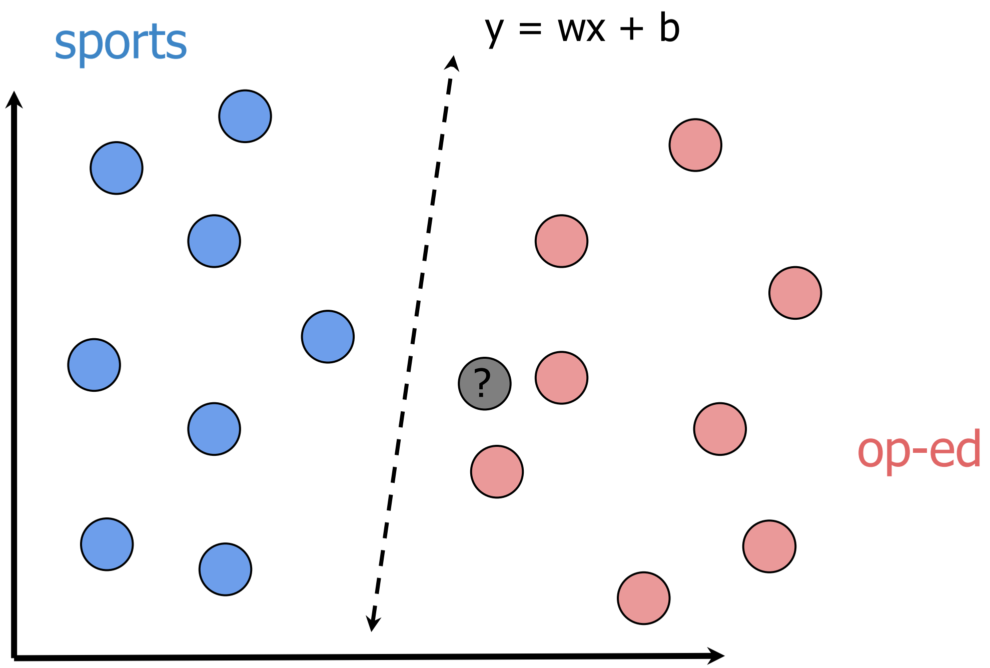
Assumptions: similar inputs have similar outputs



Draw a line

Intuition: draw a line separating classes of data

Assumptions: data is linearly separable

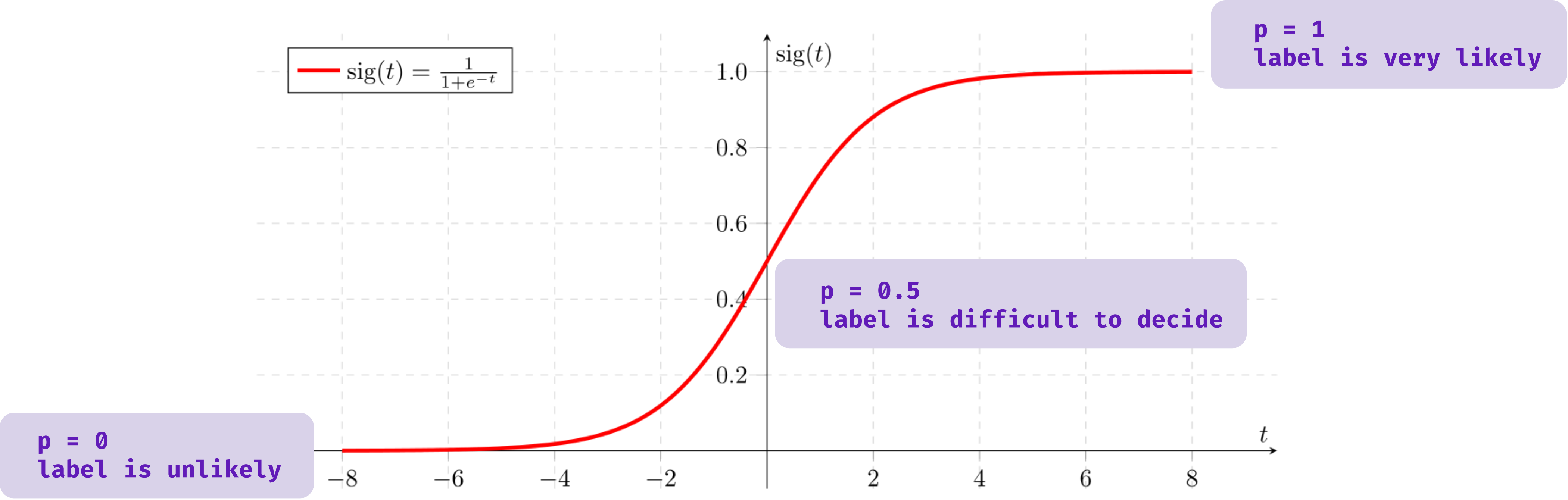


Draw a line and use it as your label probability

Intuitions

Predict the **probability** that an input belongs to a class with a sigmoid curve

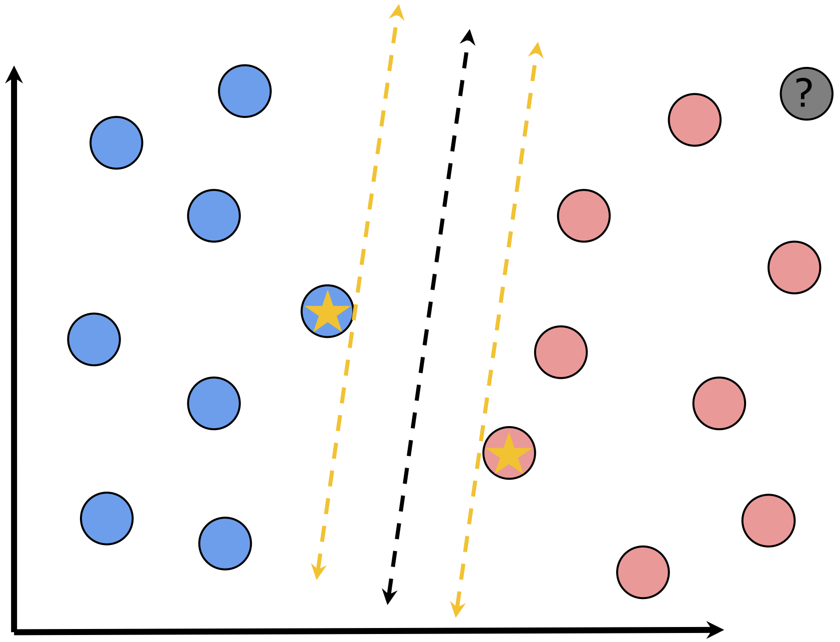
We transform or "squish" linear regression into a [0,1] range



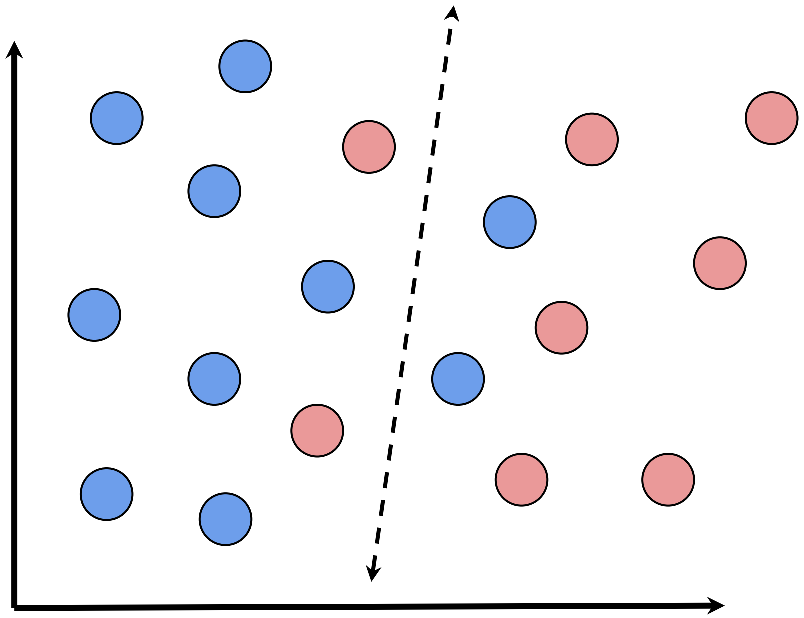
Draw a line that maximizes the distance between the groups

Intuition: find the line separating the classes that has the **maximum margin**

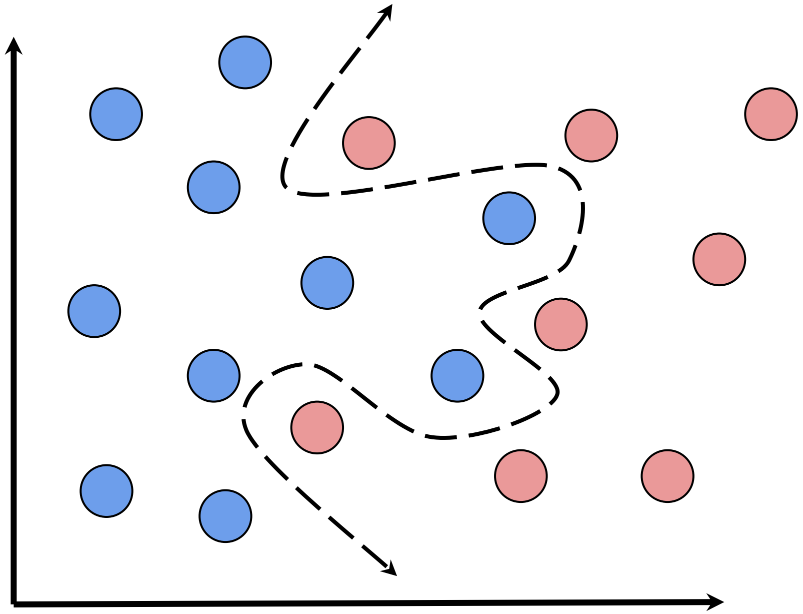
Assumption: support vectors are the most useful data points



The bias-variance tradeoff



Bias
under-fit
makes inaccurate assumptions
model is too simple



Variance
over-fit
sensitive to noise
model is too complex

Loss Function and Regularization

m = number of documents
y_i = true label
wx_i +b = predicted label

n = number of features
w = feature weight
lambda = strength of regularization

$$J(w) = \frac{1}{m}\sum_{i=0}^{m} |(wx_i + b) - y_i| + \lambda \sum_{j=0}^{n} |w_j|$$

Cost Function

Loss Function

punish the model for labeling documents incorrectly

Regularization

punish the model for being too complex

Unsupervised Methods: Examples

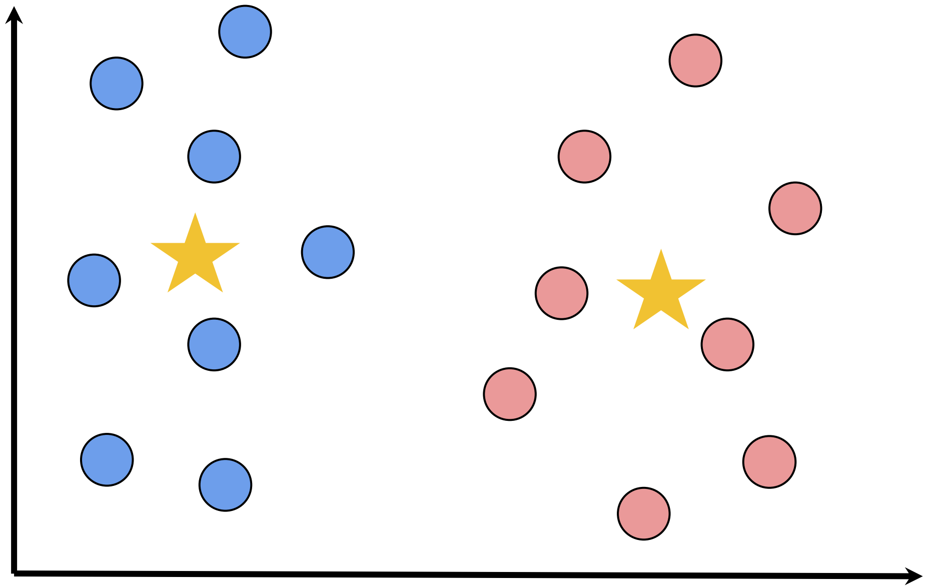
Let's take a look at some unsupervised machine learning models!

These models assume that **all you have is your data** (no labels), and you want to explore that data and see what patterns and structure you can find.

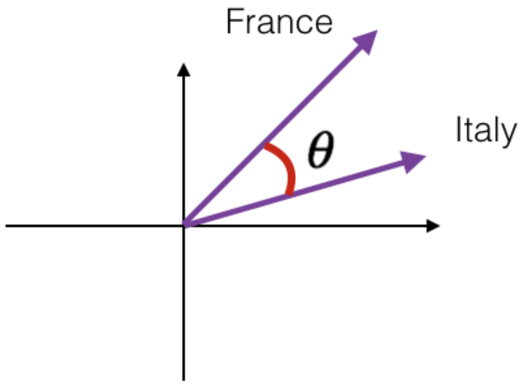
Find clusters

Intuitions:

- assign each data point to the nearest cluster centroid
- centroids should be the mean of all the points in the cluster



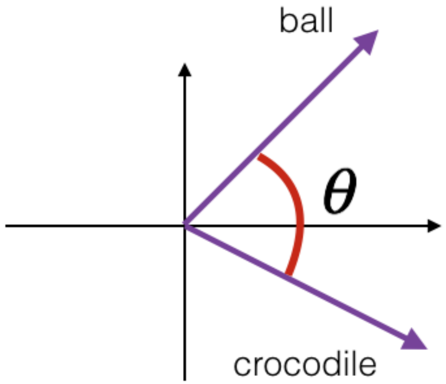
Measure relationships between words



France and Italy are quite similar

θ is close to 0°

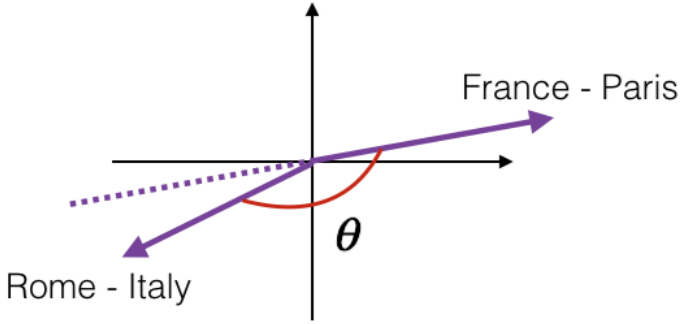
$\cos(\theta) \approx 1$



ball and crocodile are not similar

θ is close to 90°

$\cos(\theta) \approx 0$



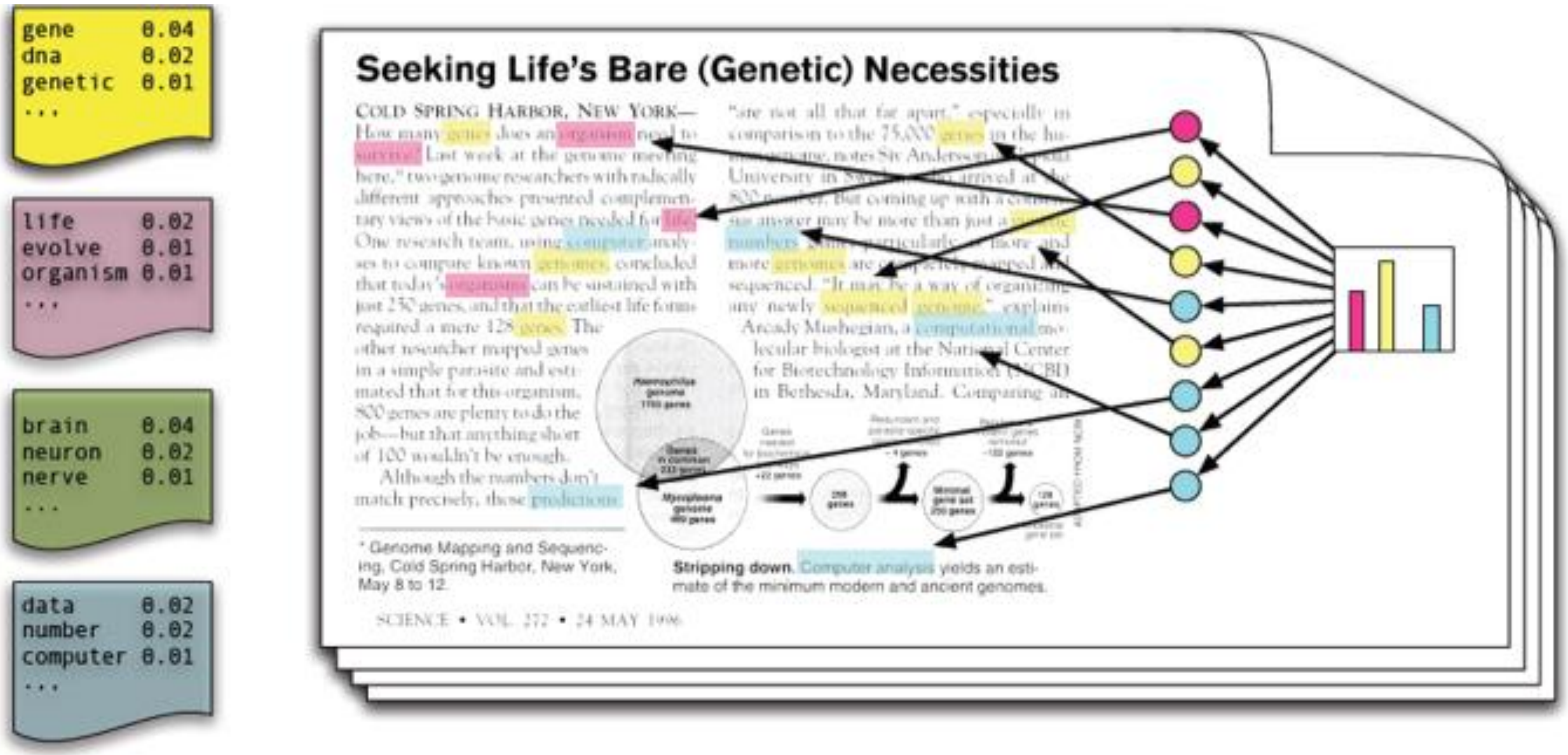
the two vectors are similar but opposite
the first one encodes (city - country)
while the second one encodes (country - city)

θ is close to 180°

$\cos(\theta) \approx -1$

https://datascience-enthusiast.com/DL/Operations_on_word_vectors.html

Discover and track topics in a dataset



Introduction to language modeling

What is a language model?

“a model that assigns a probability to sequences of words”

(Jurafsky & Martin, *Speech and Language Processing*)

Given a sequence of words, can we **predict** the next sequence of words?

We rely on some text **dataset** to estimate these probabilities.

A simple language model

P(“*question*”|“*to be or not to be, that is the*”) = ?

A simple language model

P(“*question*”|“*to be or not to be, that is the*”) = ?

Unigrams: count how many times “*question*” follows “*the*” in our dataset.

Bigrams: count how many times “*question*” follows “*is the*” in our dataset.

Trigrams: count how many times “*question*” follows “*that is the*” in our dataset

Etc.

What are “large language models”?

Also referred to as **pretrained models** and **foundation models**.

These models rely on vast amounts of pretraining data. While some performance gains come from the model architectures, a lot is coming from just the sheer amount of pretraining data.

Common pretraining sources include web scrapes, Wikipedia, Reddit, scientific publications, and books.

Examples: *ChatGPT, Claude, BERT, Galactica, BLOOM*

Ethics of large language models

Selected readings on NLP ethics

BOOKS

- *Data Feminism* by Catherine D'Ignazio & Lauren F. Klein
- *Race After Technology* by Ruha Benjamin
- *Sorting Things Out* by C. Bowker and Susan Leigh Star
- *Automating Inequality* by Virginia Eubanks

PAPERS

- Ethical Machine Learning in Health Care (Chen et al., 2021)
- Datasheets for Datasets (Gebru et al., 2021)
- The Values Encoded in Machine Learning Research (Birhane et al., 2022)
- A Survey of Race, Racism, and Anti-Racism in NLP (Field et al., 2021)
- Language (Technology) is Power: A Critical Survey of "Bias" in NLP (Blodgett et al., 2020)

Some risks to keep in mind

1. Lack of interpretability

2. Giant datasets that are very difficult to document

3. Poor representation and quality for non-English languages

4. Toxicity/bias that is baked into and even enhanced by the models

Our own sensemaking biases

Property	Human-Human Context	Human-Machine Context
Identity Construction	Sensemaking is a question about who I am as indicated by the discovery of how and what I think.	Given multiple explanations, people will internalize the one(s) that support their identity in positive ways.
Social	What I say and single out and conclude are determined by who socialized me and how I was socialized, and by the audience I anticipate will audit the conclusions I reach.	Differences in micro- and macro-social contexts affect the effectiveness of explanations.
Retrospective	To learn what I think, I look back over what I said earlier.	Providing explanations before people can reflect on the model and its predictions negatively affects sensemaking.
Enactive	I create the object to be seen and inspected when I say or do something.	The order in which explanations are seen affects how people understand a model and its predictions.
Ongoing	Understanding is spread across time and competes for attention with other ongoing projects, by which time my interests may already have changed.	The valence and magnitude of emotion caused by an interruption during the process of understanding explanations from interpretability tools change what is understood.
Focused on Extracted Cues	The 'what' that I single out and embellish is only a small portion of the original utterance, that becomes salient because of context and personal dispositions.	Highlighting different parts of explanations can lead to varying understanding of the underlying data and model.
Plausibility over Accuracy	I need to know enough about what I think to get on with my projects, but no more, which means sufficiency and plausibility take precedence over accuracy.	Given plausible explanations for a prediction, people are not inclined to search for the accurate one amongst these.

Table 1: An overview of the seven properties of sensemaking, their description in the human-human context, and our proposed claims for the human-machine context grounded in each property.

"Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning" (Kaur et al., 2020)

"Sensible AI: Re-imagining Interpretability and Explainability using Sensemaking Theory" (Kaur et al., 2022)

From model to action

Prompting and finetuning

Start with a base model trained on a bunch of random data.

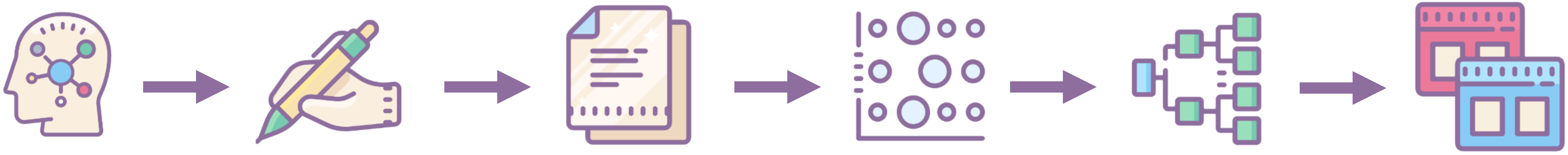
Add task-specific heads that produce specific output patterns

You can do this via

- **Finetuning**: adjusting the model to a small amount of target data
- **Prompting**: provide the first part of your output as a "hint" of where to go

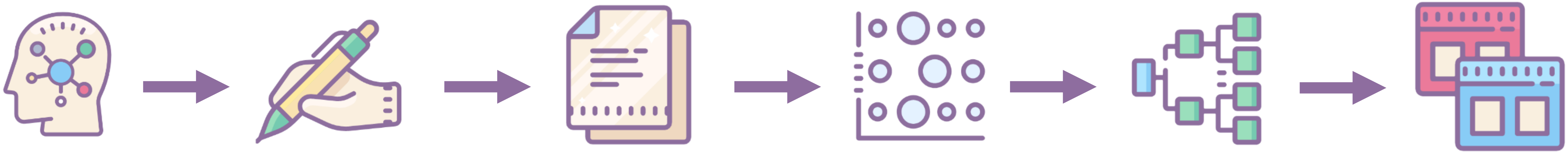
NLP, ML, Industry

Downstream



NLP, ML, Industry

Downstream



Upstream

Computational Social Science, Digital Humanities

Quantitative + qualitative methods

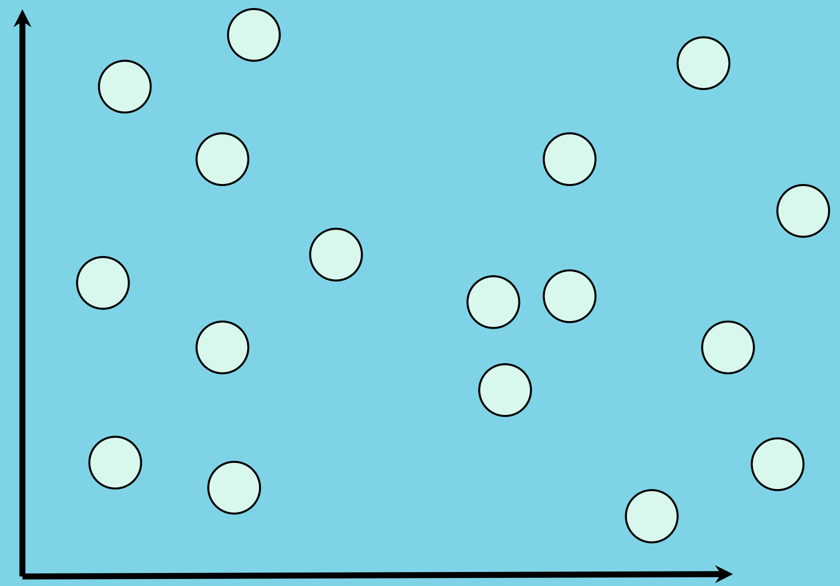
Quantitative methods like those used in NLP are great at finding patterns and averages.

But they can also be used to find outliers and interesting cases.

And those results can lead us back to specific stories and the direct words of people seeking care.

Thank you!

I’ll see you all again later today when we’ll do an interactive exercise where you’ll get to try out a language model for yourself!



8	32	0	0	3	99
74	0	5	45	0	0
3	26	0	5	0	92
0	0	16	0	64	23

MATERNAL HEALTH EQUITY WORKSHOP:
FROM STORY TO DATA TO ACTION
MAY 18, 2023

From Data to Research:
NLP & Maternal Health Use Cases

Association of
American Medical Colleges

Meet the Speakers

Allan Fong, MS
Senior Research Scientist and Data Scientist, MedStar Health

Angela D. Thomas, DrPH, MPH, MBA
Vice President, Healthcare Delivery Research, MedStar Health

Mark Clapp, MD, MPH
Maternal-Fetal Medicine Specialist, Physician Investigator, Massachusetts General Hospital

Anna Wexler, PhD
Assistant Professor of Medical Ethics and Health Policy, University of Pennsylvania Perelman School of Medicine

Amulya Yadav, PhD
Assistant Professor, PNC Technologies Career Development and Associate Director, Center for Socially Responsible Artificial Intelligence, Penn State University

MedStar Health

It's how we treat people.

May 18, 2023

The Role of NLP in Addressing Maternal Harm through a Patient Safety and Equity Lens

Angela D. Thomas, DrPH, Vice President, Healthcare Delivery Research, MedStar Health Research Institute

Allan Fong, MS, Senior Research and Data Scientist, MedStar Health Research Institute

Agenda

1. Safe Babies Safe Moms Maternal Taxonomy Study
   - Background
   - Disparities in Maternal Harm
   - Addressing Disparities through a Patient Safety & Health Equity Lens
   - Aims
2. Deeper Dive on NLP Aim
3. Future Directions

SBSM Maternal Taxonomy Supplemental Study

D.C. Safe Babies Safe Moms.

MedStar Health

MedStar Health Research Institute

Note on Inclusivity

Not all birthing individuals identify as women or mothers. Throughout this presentation, you may hear the terms women and mothers to most accurately reflect the current state of the literature, but our approach is meant to be inclusive of all birthing individuals.

About Safe Babies Safe Moms

D.C. Safe Babies Safe Moms.

A. JAMES & ALICE B. CLARK FOUNDATION

- April 2020 – March 2025
- $30M initiative
- Largest philanthropic donation in MedStar history

Reduce disparities in maternal & infant mortality in Washington, D.C.

Key Literature Facts

- Maternal harm is a major crisis that disproportionately affects Black women
- Education is not a protective factor
- ~80% of maternal deaths and ~90% of severe maternal morbidity events are preventable
- Most frequent preventable factors are provider-related and/or system-related

Figure 1: The Safety Spectrum

Figure 2: Patient Safety Taxonomy Adapted from the MERP Index for Categorizing Medication Errors

Our Proposed Study

Current Literature

Category	Description
A	Unsafe Condition (Non-Event)
B1	Near miss-No Harm - Didn't Reach Patient/Caught by Chance
B2	Near miss-No Harm - Didn't Reach Patient because of Active Recovery by Caregivers
C	No Harm - Reached Patient - No Monitoring Required
D	No Harm - Reached Patient - Monitoring Required
E	Harm - Temporary Harm - Intervention Needed
F	Harm - Temporary Harm - Hospitalization Needed
G	Harm- Permanent Harm
H	Harm-Permanent Harm - Intervention Required to Sustain Life
I	Death

More Key Literature Facts

- Racism, discrimination, and implicit bias contribute to adverse health outcomes that include adverse maternal outcomes
- Giving Voice to Mothers study found that women of color are more likely to experience mistreatment during the perinatal period
  - Shouted at, scolded, threatened, ignored, or receiving no response to requests for help
- Stereotypes stigmatizing Black motherhood including assumptions about being single, low income, and having multiple children

What doesn’t get documented…

breath. Due to her history of pulmonary embolisms (Williams **underwent emergency treatment** for a life-threatening embolism in 2011), the tennis star quickly alerted a nurse about her symptoms.

But the response wasn’t what she expected. Vogue writer Rob Haskell explains:

> She walked out of the hospital room so her mother wouldn’t worry and told the nearest nurse, between gasps, that she needed a CT scan with contrast and IV heparin (a blood thinner) right away. The nurse thought her pain medicine might be making her confused. But Serena insisted, and soon enough a doctor was performing an ultrasound of her legs. “I was like, a Doppler? I told you, I need a CT scan and a heparin drip,” she remembers telling the team. The ultrasound revealed nothing, so they sent her for the CT, and sure enough, several small blood clots had settled in her lungs. Minutes later she was on the drip. “I was like, listen to Dr. Williams!”

- If she died…”pre-existing conditions” (not bias)

Story after story after story…

There was the new mother in Nebraska with a history of hypertension who couldn't get her doctors to believe she was having a heart attack until she had another one. The young Florida mother-to-be whose breathing problems were blamed on obesity when in fact her lungs were filling with fluid and her heart was failing. The Arizona mother whose anesthesiologist assumed she smoked marijuana because of the way she did her hair. The Chicago-area businesswoman with a high-risk pregnancy who was so upset at her doctor's attitude that she changed OB/GYNs in her seventh month, only to suffer a fatal postpartum stroke.

npr

ProPublica

Kira Johnson

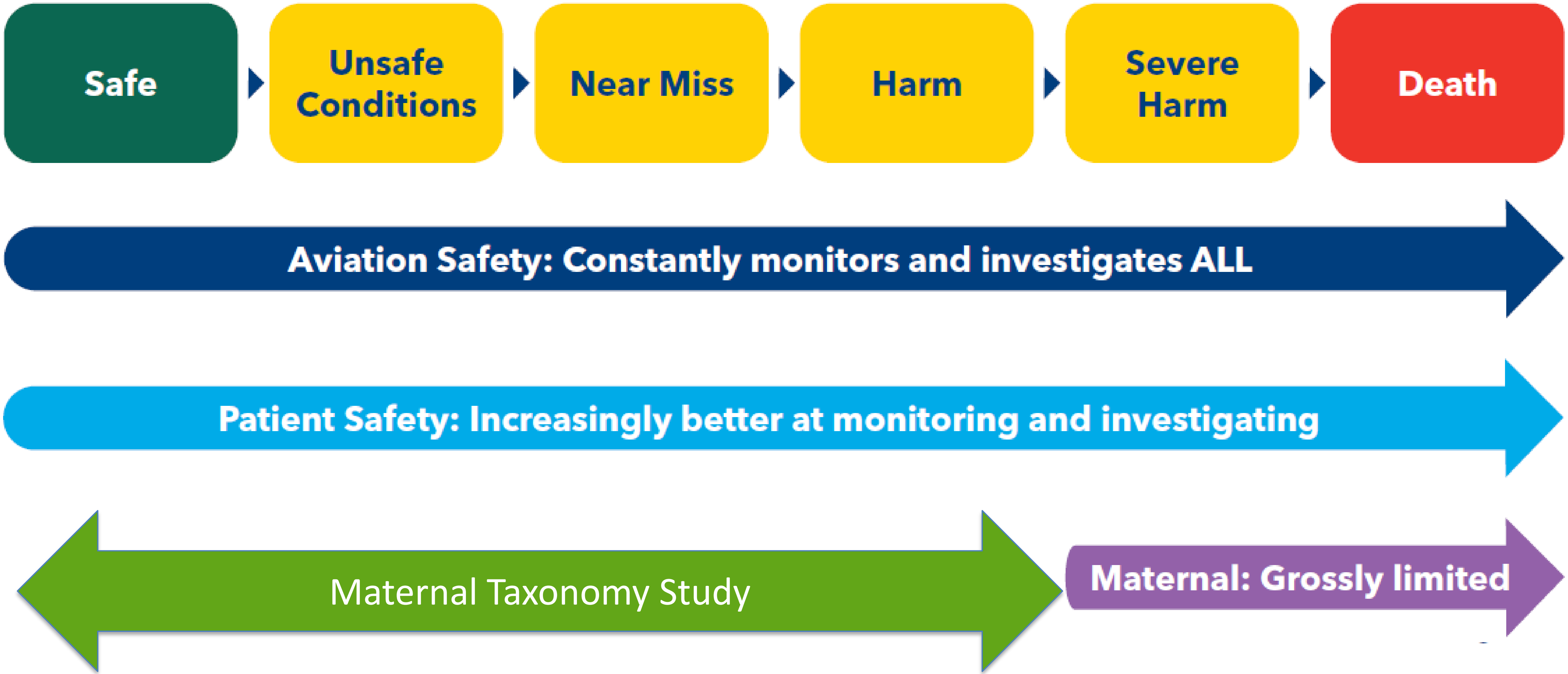
Charles Johnson (her widower) and their sons

Including the birthing individual's voice can uncover the role of racism, discrimination, and implicit bias across the full spectrum of maternal harm.

Opportunities to Systematically Remove Bias

Key Innovations

- Expertise in using clinical informatics techniques to identify "signals" in the EHR of unsafe care delivery
  - Identify "signals" in the electronic record for unsafe maternal care delivery
- Data science expertise in analyzing clinic notes in the EHR for tone and sentiment in preventable patient safety events
  - Apply to clinic notes for maternal care to uncover the role of differential treatment and biases


Safe
Unsafe Conditions
Near Miss
Harm
Severe Harm
Death
Aviation Safety: Constantly monitors and investigates ALL
Patient Safety: Increasingly better at monitoring and investigating
Maternal Taxonomy Study
Maternal: Grossly limited

Aim 1 (Phase 1)

1. Identifying common themes and signals that contribute to unsafe conditions, hazards, near misses, and other maternal injuries by:
   - Incorporating birthing individual voice – quantitative and qualitative interviews
   - Analysis of patient complaints
   - Common signals in EHR
   - Voluntary occurrence reporting system analysis
   - NLP Analysis of EHR Notes

Aim 2 (Phase 2)

2. Developing a full maternal safety spectrum taxonomy by identifying how the common themes and signals contribute to a progression from:
   - Safe to unsafe conditions
   - Unsafe conditions to hazards
   - Hazards to near misses
   - Near misses to maternal harm
   - Maternal harm to severe maternal morbidity
   - Severe maternal morbidity to maternal mortality

Aim 3 (Phase 2)

3. **Developing a toolkit** for ongoing surveillance of common themes and signals, with mitigation strategies when they are identified.

Local & National Interdisciplinary Expertise

- Maternal qualitative survey experts
- Health equity experts
- Obstetrics experts
- Midwifery experts
- Human factors experts
- Patient safety experts
- Data science experts
- Informatics experts
- Biostatistics experts

Aim 1 (Phase 1)

1. Identifying common themes and signals that contribute to unsafe conditions, hazards, near misses, and other maternal injuries by:
   - Incorporating birthing individual voice – quantitative and qualitative interviews
   - Analysis of patient complaints
   - Common signals in EHR
   - Voluntary occurrence reporting system analysis
   - NLP Analysis of EHR Notes

A Deeper Dive

Using NLP - Biases in EHR Notes

MedStar Health

MedStar Health Research Institute

Method - Motivation

RESEARCH ARTICLE | HEALTH EQUITY

HEALTH AFFAIRS > VOL. 41, NO. 2: RACISM & HEALTH

Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record

Michael Sun, Tomasz Oliwa, Monica E. Peek, and Elizabeth L. Tung

AFFILIATIONS

PUBLISHED: JANUARY 19, 2022 Open Access https://doi.org/10.1377/hlthaff.2021.01423

SECTIONS VIEW ARTICLE PERMISSIONS SHARE TOOLS

Abstract

Little is known about how racism and bias may be communicated in the medical record. This study used machine learning to analyze electronic health records (EHRs) from an urban academic medical center and to investigate whether providers' use of negative patient descriptors varied by patient race or ethnicity. We analyzed a sample of 40,113 history and

- January 2019–October 2020
- 18,459 patients
- 40,113 history and physical notes
- Sentences containing a negative descriptor
- Compared with White patients, Black patients had 2.54 times the odds of having at least one negative descriptor in the history and physical notes

Method

- Data Source
  - Jan. 1, 2016 to March 31, 2020
  - History and physical, review of system, history of present illness, physical examination, triage notes

Study cohort

↓

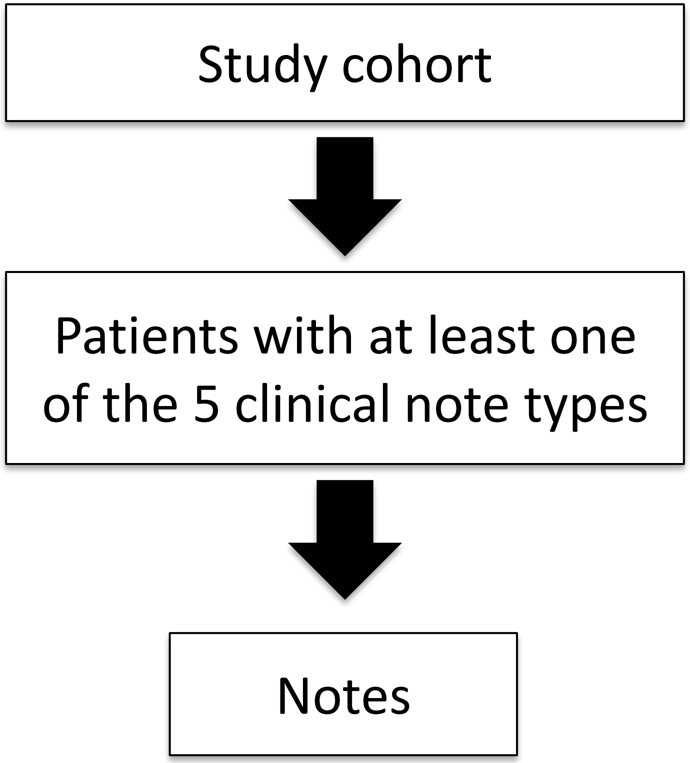
Patients with at least one of the 5 clinical note types

↓

Notes

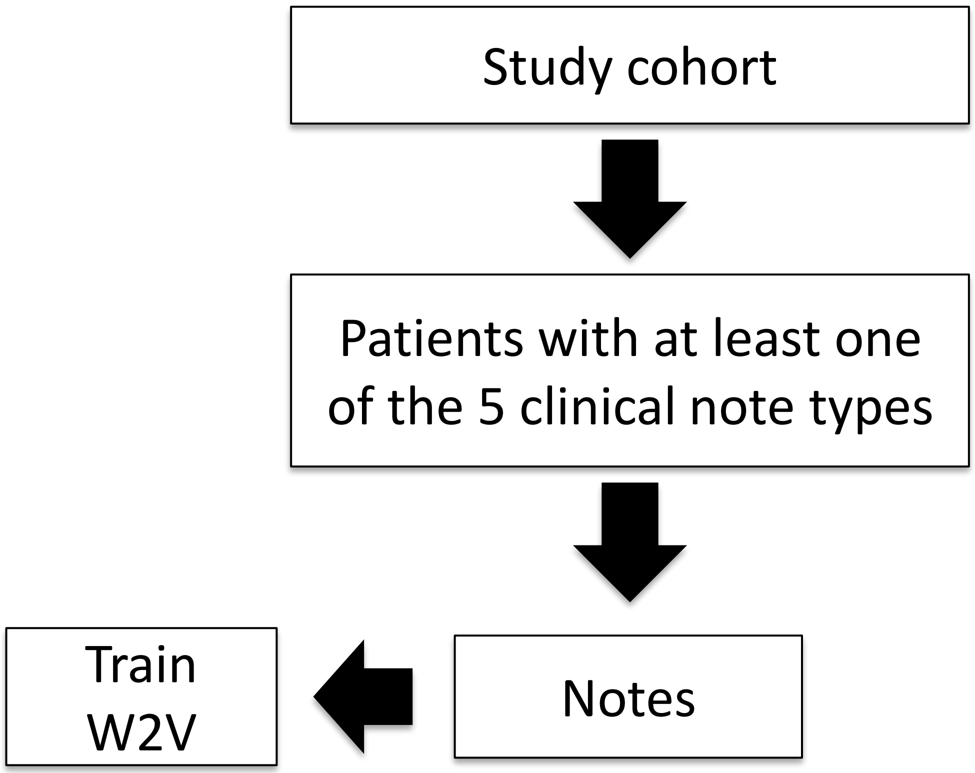
Method

- Data Source
  - Jan. 1, 2016 to March 31, 2020
  - History and physical, review of system, history of present illness, physical examination, triage notes
- Negative Descriptors
  - Nonadherent, aggressive, agitated, angry, challenging, combative, noncompliant, confront, noncooperative, defensive, exaggerate, hysterical, unpleasant, refuse, and resist
  - Descriptor and root
  - Query expansion using W2V (Why use NLP)



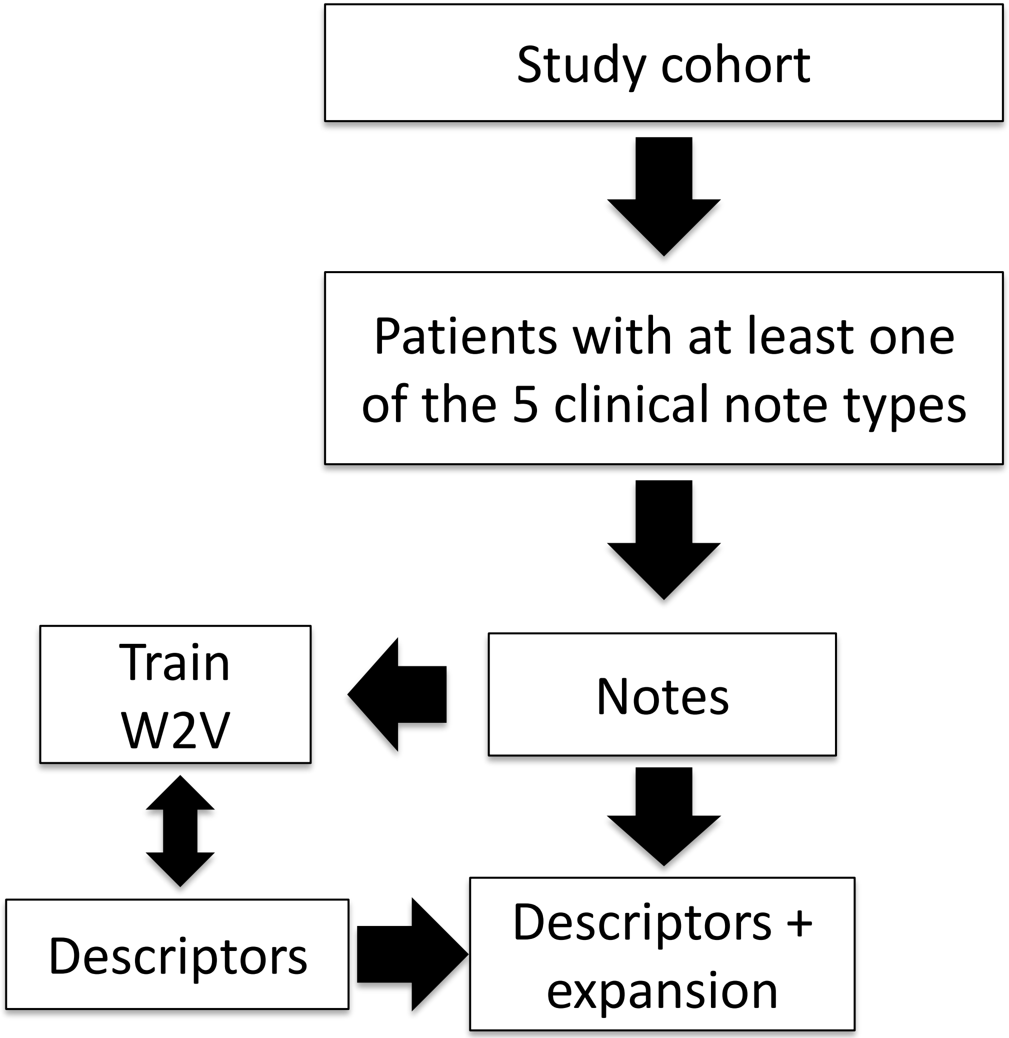
Method

- Data Source
  - Jan. 1, 2016 to March 31, 2020
  - History and physical, review of system, history of present illness, physical examination, triage notes
- Negative Descriptors
  - Nonadherent, aggressive, agitated, angry, challenging, combative, noncompliant, confront, noncooperative, defensive, exaggerate, hysterical, unpleasant, refuse, and resist
  - Descriptor and root
  - Query expansion using W2V (Why use NLP)



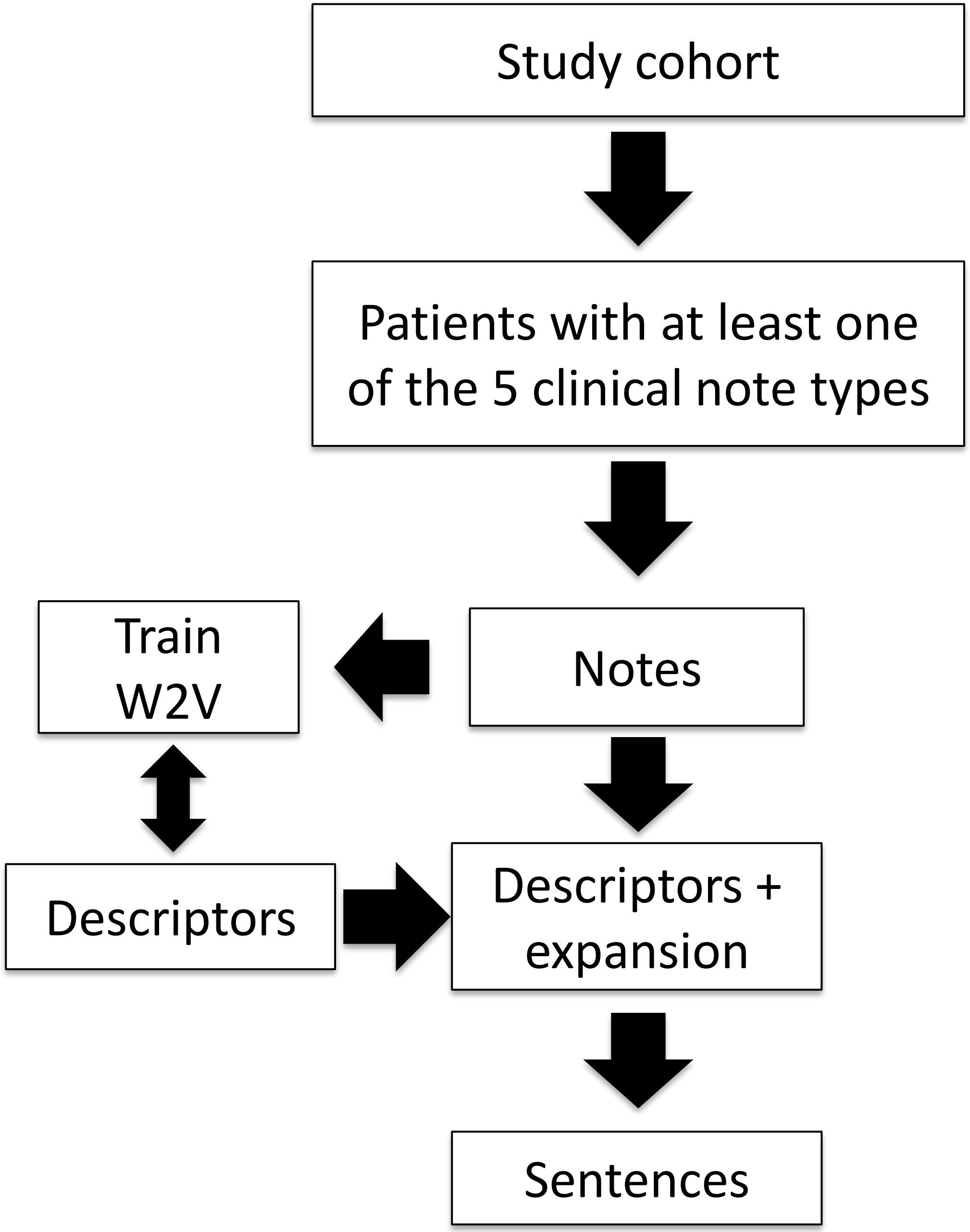
Method

- Data Source
  - Jan. 1, 2016 to March 31, 2020
  - History and physical, review of system, history of present illness, physical examination, triage notes
- Negative Descriptors
  - Nonadherent, aggressive, agitated, angry, challenging, combative, noncompliant, confront, noncooperative, defensive, exaggerate, hysterical, unpleasant, refuse, and resist
  - Descriptor and root
  - Query expansion using W2V (Why use NLP)



Method

- Data Source
  - Jan. 1, 2016 to March 31, 2020
  - History and physical, review of system, history of present illness, physical examination, triage notes
- Negative Descriptors
  - Nonadherent, aggressive, agitated, angry, challenging, combative, noncompliant, confront, noncooperative, defensive, exaggerate, hysterical, unpleasant, refuse, and resist
  - Descriptor and root
  - Query expansion using W2V (Why use NLP)

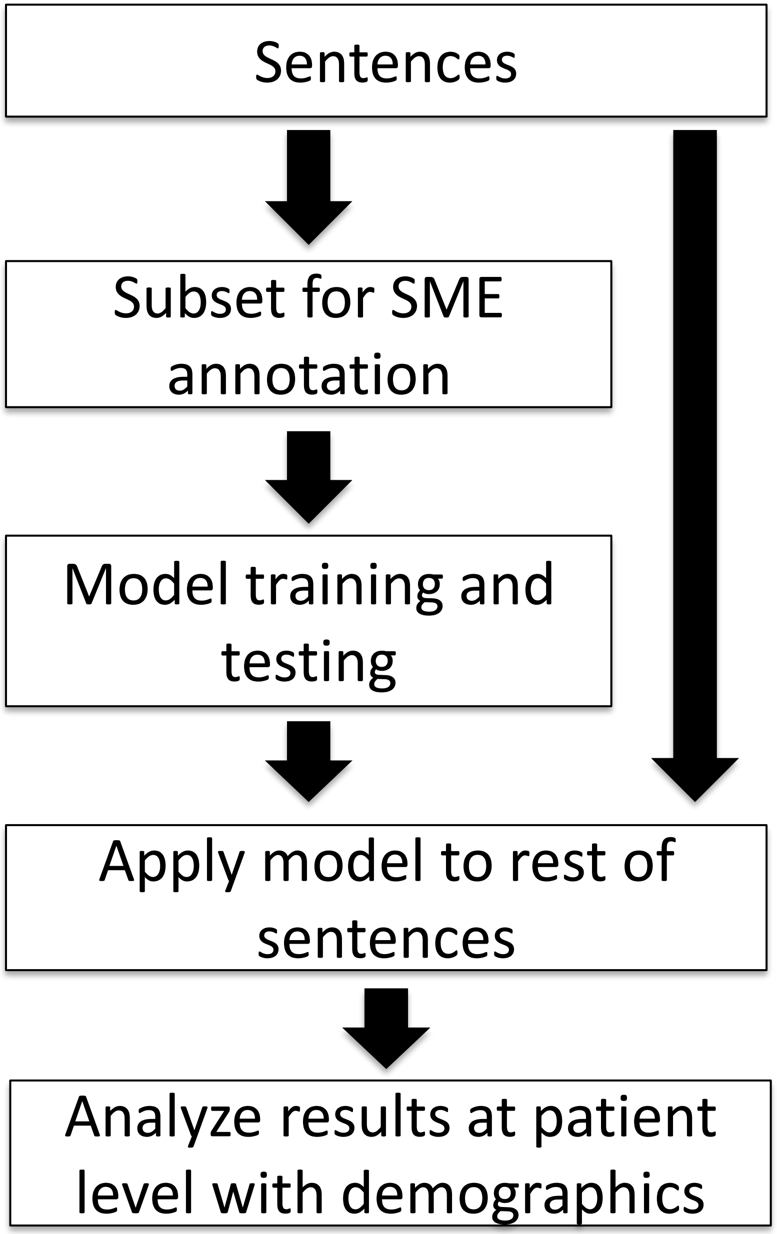


Method

- ‘aggressive’
  - “displaying irritable mood and aggression towards staff…”
- angry
  - “she felt as if the previous nurse was laughing at her and this made her very angry…”
- exaggerated
  - “her exaggerated blame on the hospital staff…”
- defensive
  - “she became defensive and upset and said it was because of…”

Method

- Annotations
- Train model (Why use NLP)
- Applied model to unseen data
- Analyze results at patient level with demographics



Model Results

Performance metrics of the negative language classifiers

Metrics	Train Data	Test Data
Area Under Receiver Operating Characteristics Curve	0.99	0.89
Sensitivity	0.81	0.61
Specificity	0.99	0.99
Positive Predictive Value	0.91	0.95
Negative Predictive Value	0.98	0.97
Accuracy	0.98	0.96
F-1 Score	0.86	0.74
Area Under Precision-Recall Curve	0.94	0.8

Discussion

- Analysis is still underway
- Key differences align with previous literature on documentation bias
- Why use NLP
  - Data driven way to tailor search terms
    - Our healthcare system specific
  - Too many sentences to manually review
    - Limited time and resources

Future Directions

NLP & Operations

Discussion - NLP

- Negative language/racial bias
- Positive language
  - ‘a pleasant 26 yo female patient’
- Self-awareness
  - Changes in documentation tone (ie. between patients, open notes access, burn-out)

Discussion - Operations

- Aim 4 – Pilot the toolkit
- Ongoing surveillance efforts for accountability and change
  - Ongoing capture of the birthing individuals' voices
  - Ongoing analysis of patient complaints
  - Ongoing surveillance of signals in the EHR
  - Ongoing analysis of voluntary occurrence reporting system analysis
  - Ongoing analysis of EHR Notes
- Opportunities for automation using NLP?

Thank you

It's how we treat people.

MedStar Health

NLP for Risk Prediction in Maternity Care

Mark Clapp, MD MPH

Massachusetts General Hospital

Harvard Medical School

MASSACHUSETTS GENERAL HOSPITAL
OBSTETRICS & GYNECOLOGY

HARVARD MEDICAL SCHOOL

mark.clapp@mgh.harvard.edu @markaclapp

PROPUBLICA

The Last Person You'd Expect to Die in Childbirth

Why America's Black Mothers and Babies Are in a Life-or-Death Crisis

The answer to the disparity in death rates has everything to do with the lived experience of being a black woman in America.

By LINDA

CBS NEWS / *August 5, 2018, 10:06 AM*

Maternal mortality: An American crisis

Nearly Dying In Childbirth: Why Preventable Complications Are Growing In U.S.

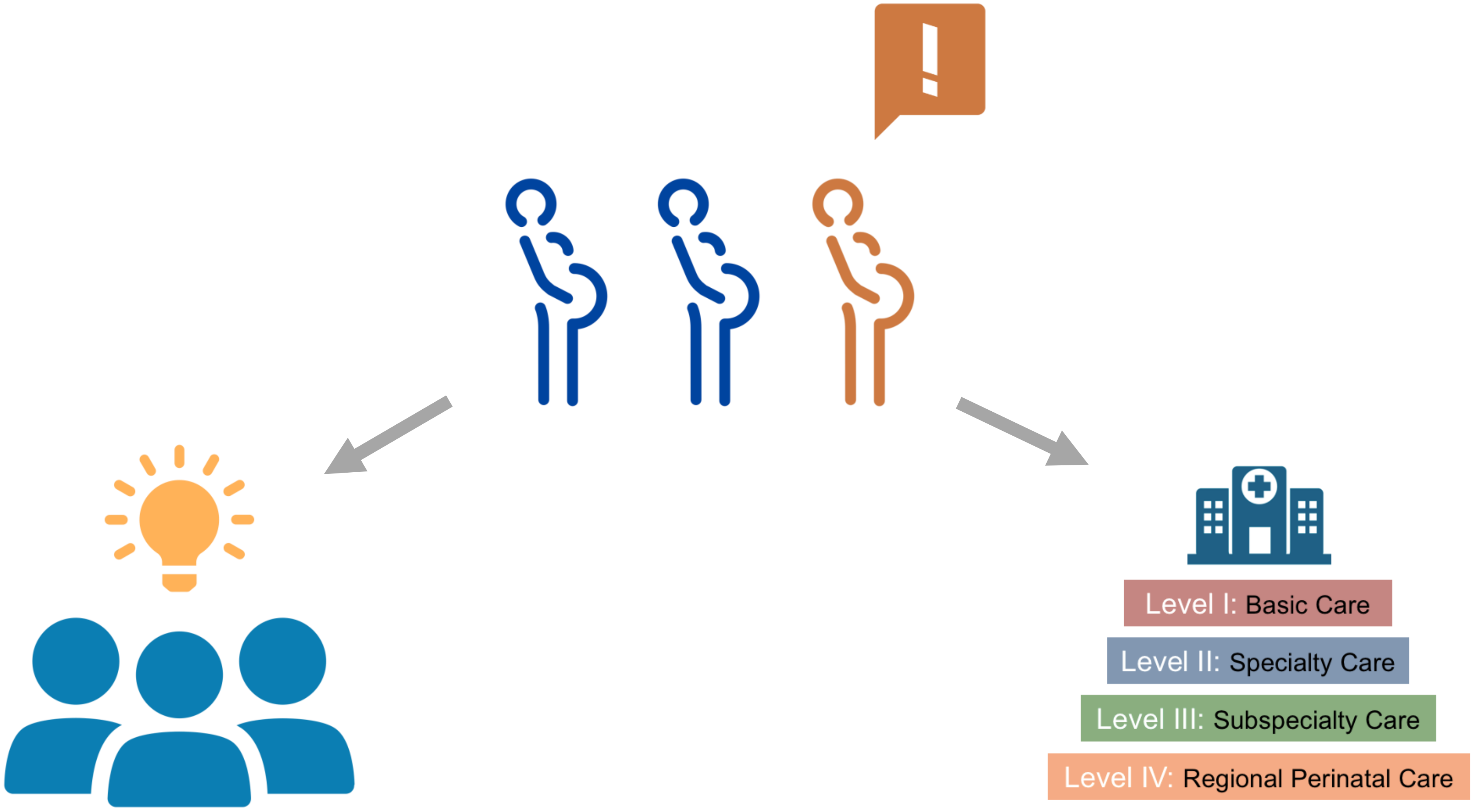
December 22, 2017 · 12:17 PM ET

KATHERINE ELLISON, PROPUBLICA

NINA MARTIN, PROPUBLICA

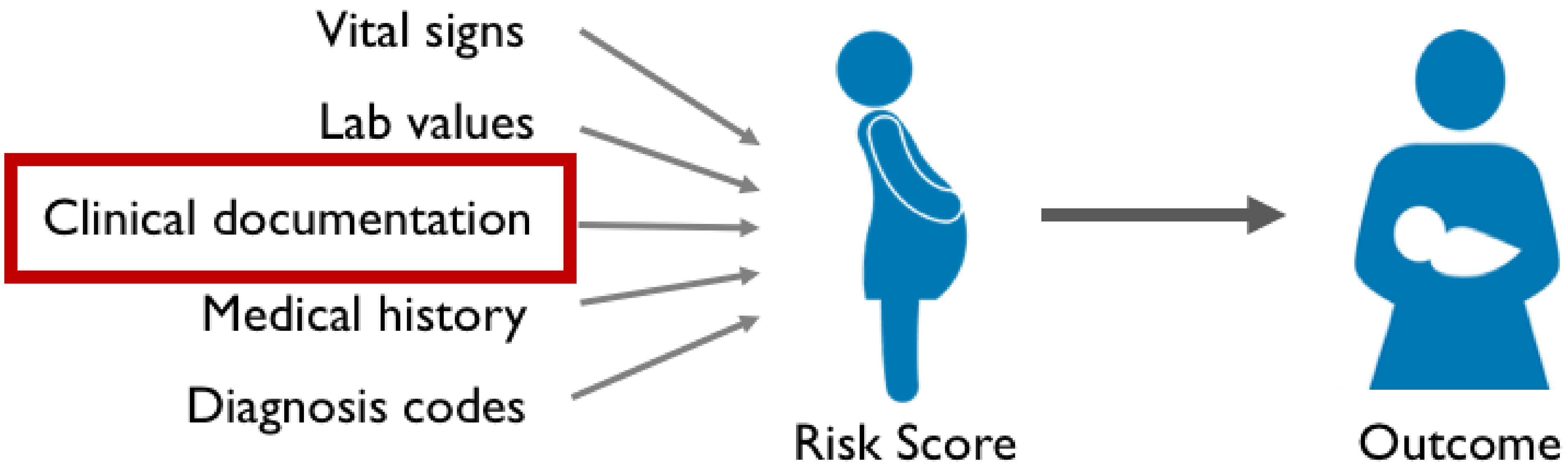
NYT; 4/11/18. ProPublica; 5/12/17. CBS News; 8/3/18. NPR; 12/22/17.

Goals of Risk Prediction

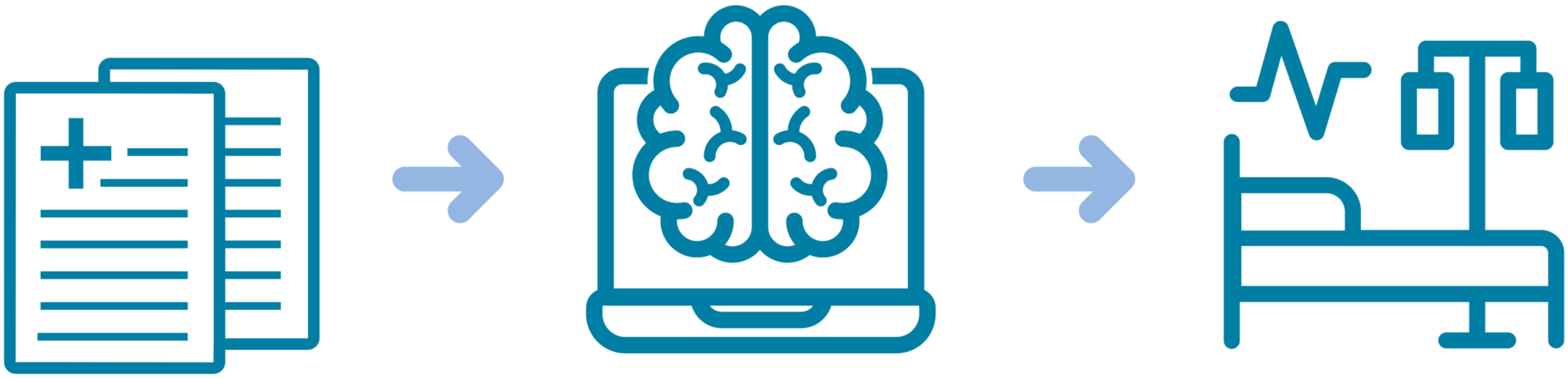


Goals of a Clinical Risk Tool

Utilize EHR information to develop methods for personalized maternal risk stratification in obstetrics



Natural Language Processing



NLP in its Simplest Form

Ms. Smith presents for induction of labor for pre-eclampsia.

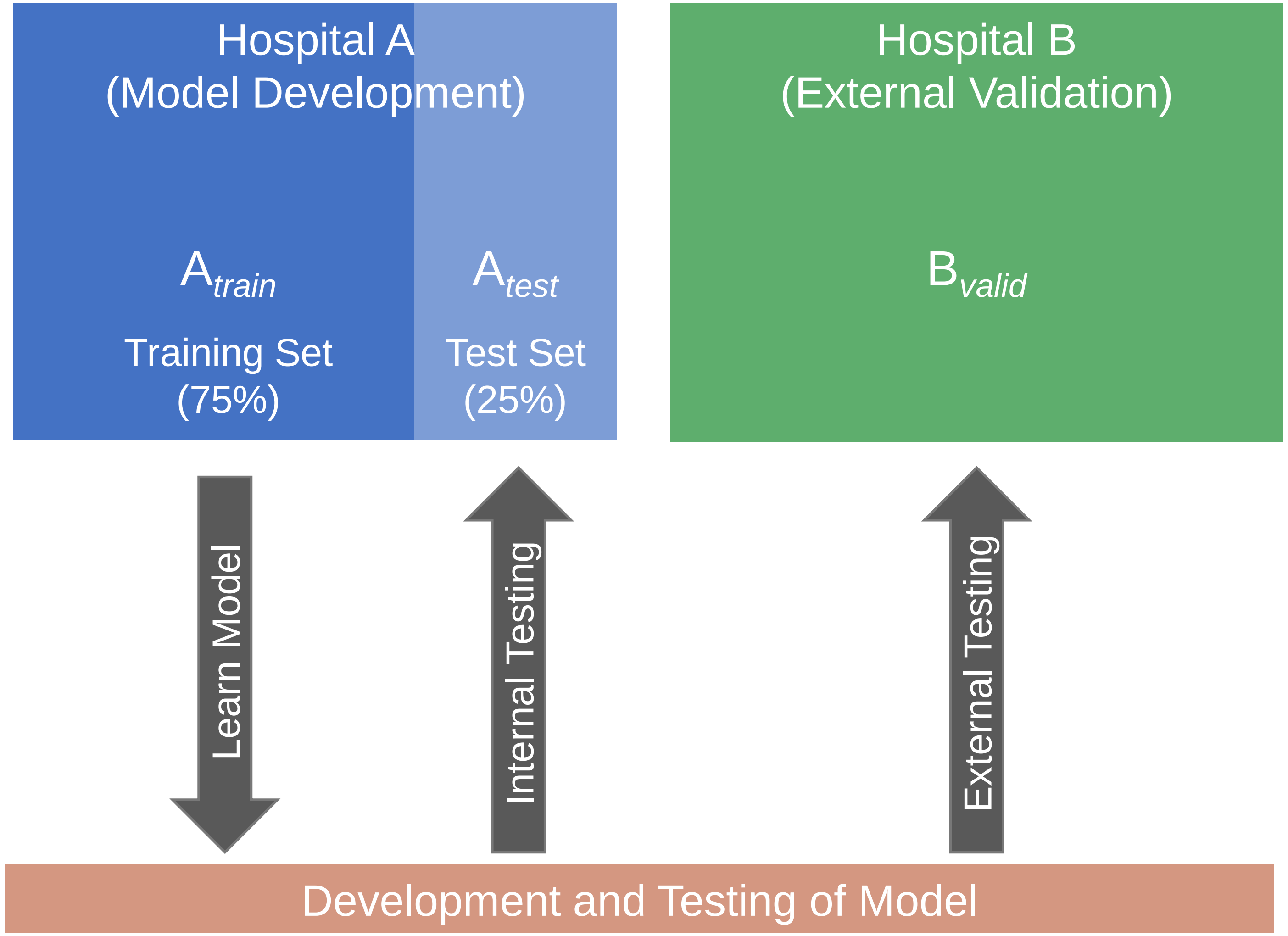
↓

ms smith presents induction labor preeclampsia

↓

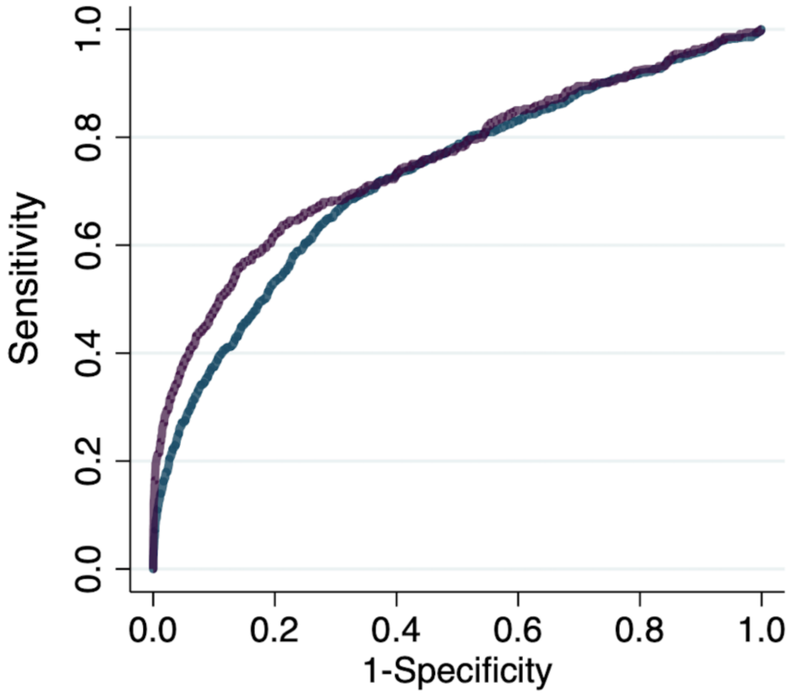
MRN#	ms	smith	presents	induction	labor	preeclampsia
123456	1	1	1	1	1	1

Model Derivation and Testing



Clapp et al. AJOG. 2022.

Model Performance



Area Under the Curve (AUC) values

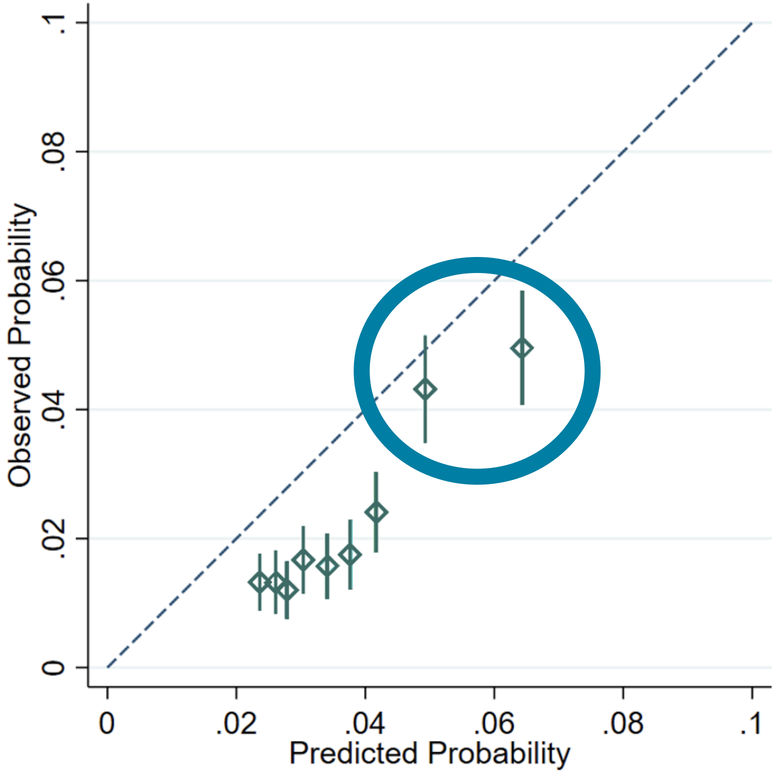
0.72 for severe morbidity

0.76 for severe morbidity (excluding transfusion)

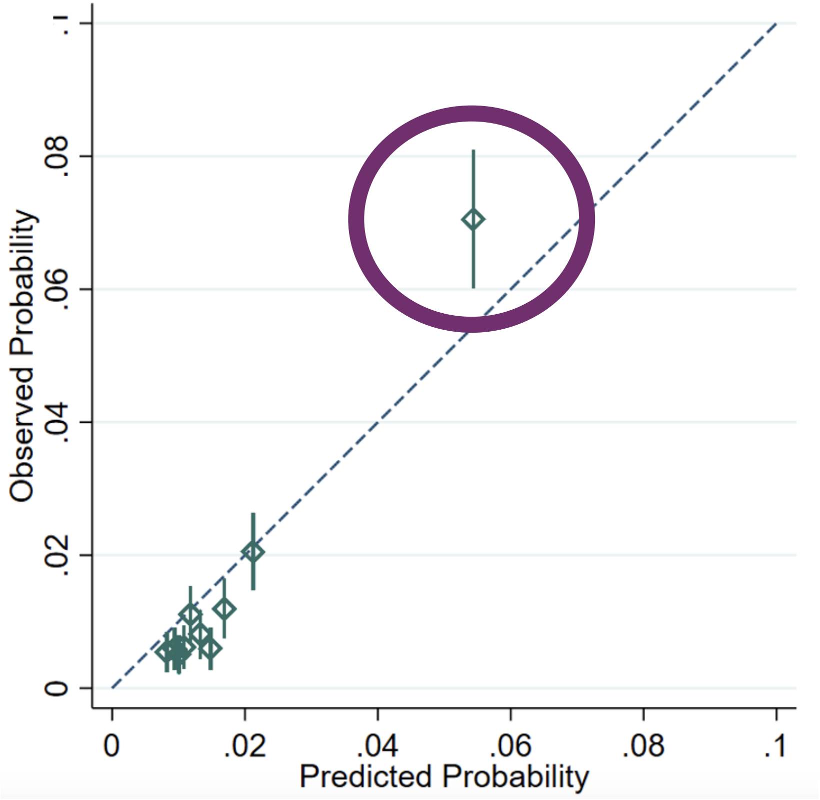
Clapp et al. AJOG. 2022.

Model Performance

Severe Morbidity

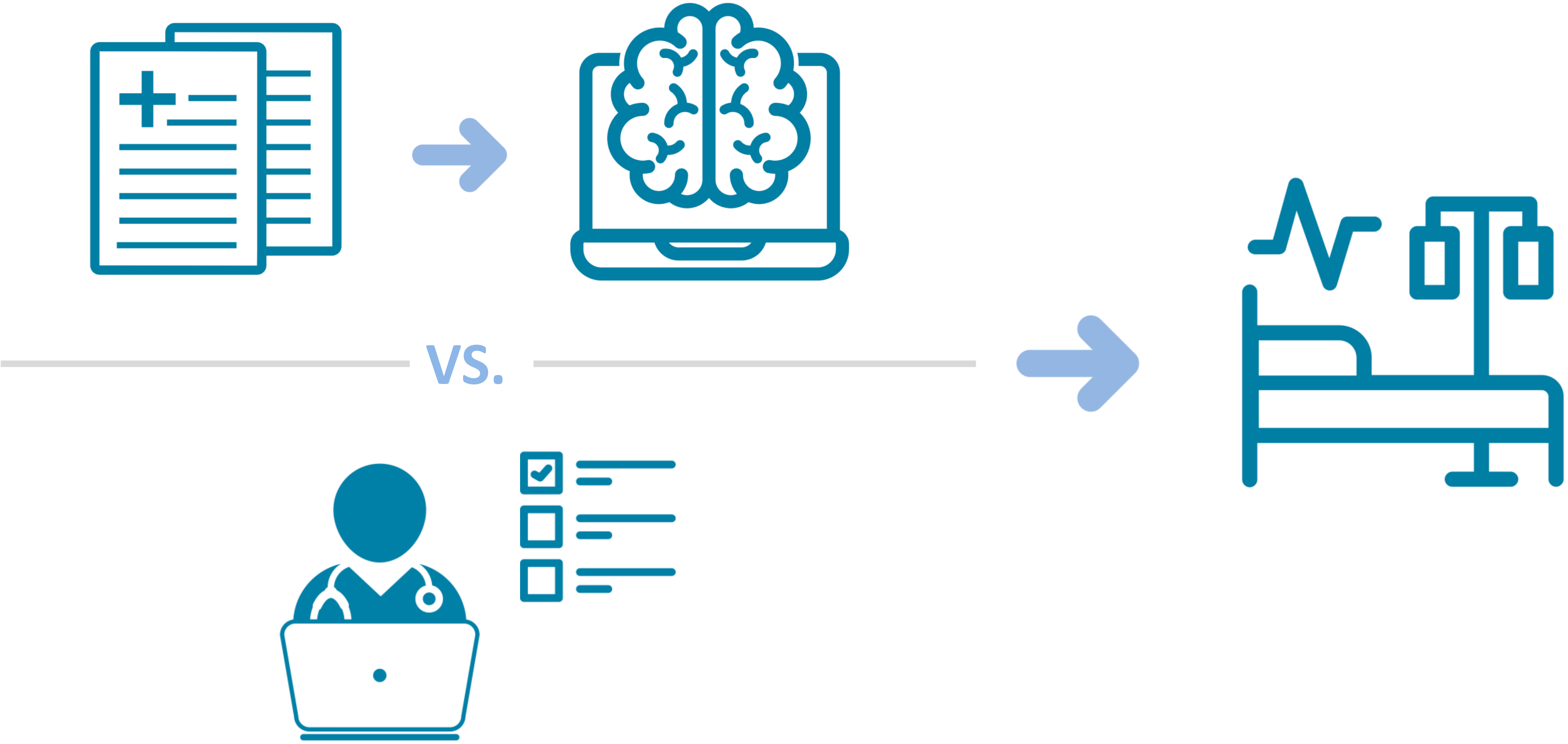


Non-Transfusion Severe Morbidity

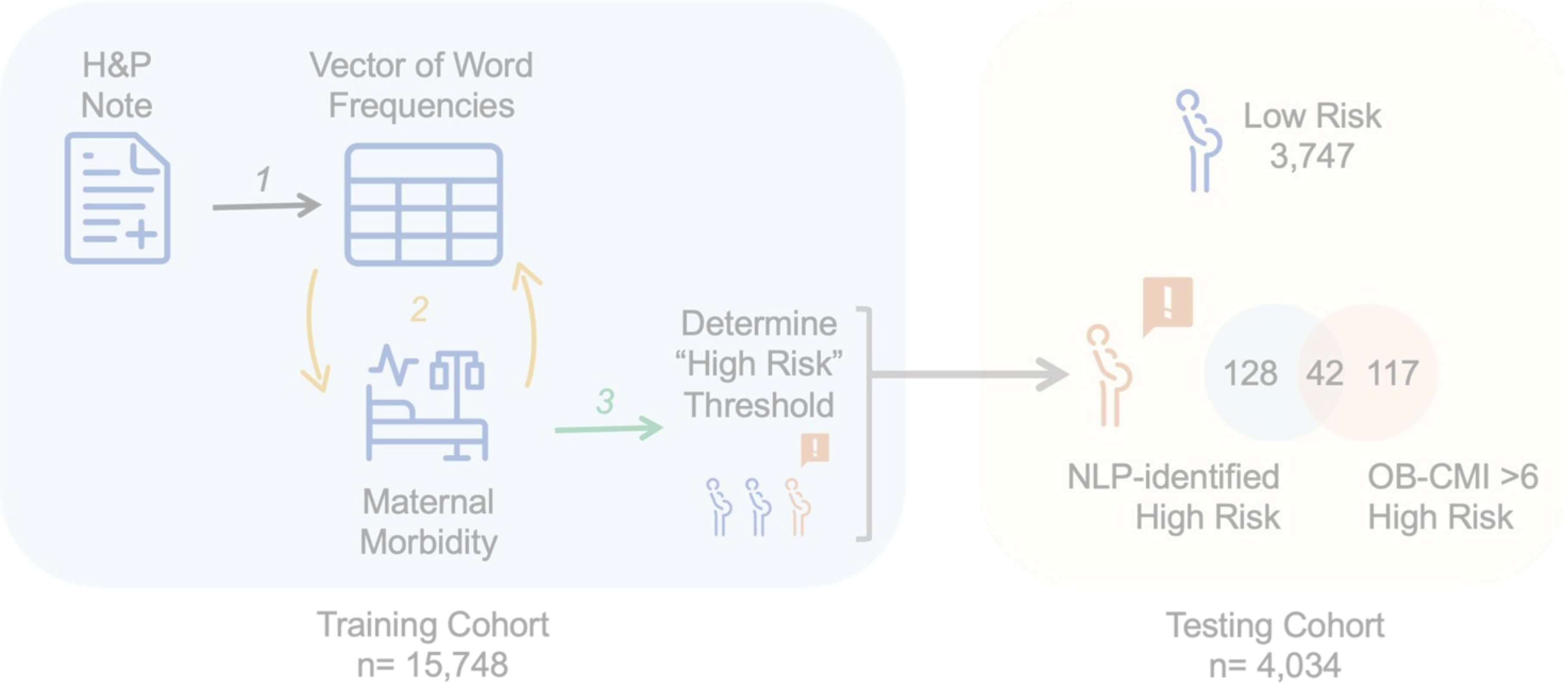


Clapp et al. AJOG. 2022.

NLP Model vs. Manual Risk Assignment



NLP Model vs. Manual Risk Assignment



Clapp et al. JAMA Network Open. Oct 2022.

Model Performance

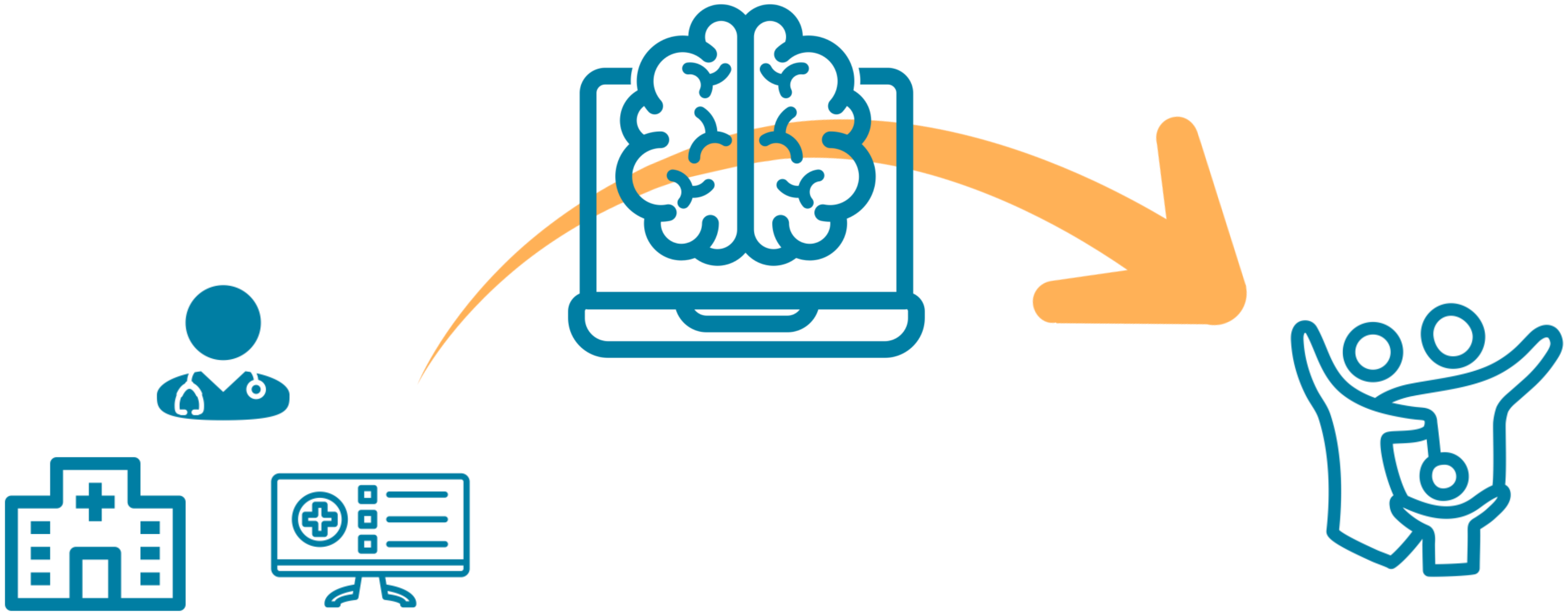
Model Characteristic	SMM OB-CMI	NLP	OB-CMI+ NLP
Screen Positive Rate	3.9%	4.2%	7.1%
Sensitivity	24.4%	28.7%	37.4%
Positive Predictive Value	17.6%	19.4%	15.0%
Specificity	96.7%	96.5%	93.8%
Negative Predictive Value	97.8%	97.9%	98.1%

Clapp et al. JAMA Network Open. Oct 2022.

Other Findings

- Similar performance when using more advanced NLP modeling techniques
  - Bi- and tri-gram models
  - Including negation
  - Using pretrained dictionaries
  - Various machine learning models
- Similar results when predicting postpartum hemorrhage, one of the most common pregnancy complications

Next Steps



Prospective application in clinical practice

Challenges for NLP Applications

Translation to the bedside

- End-user buy-in
- Adapting clinical workflows
- Facile health IT infrastructure

Ensuring performance and equity in application

- Challenges in inequities in underlying data to derive models
- Variable model performance among population subgroups
- Uncertain generalizability
- Availability of model/technology to all populations

Special Thanks

Funders

- AAOGF/ABOG Research Scholar Award
- SMFM Bridge Award

Collaborators

- Anjali Kaimal, MD MAS
- Tom McCoy, MD
- Roy Perlis, MD MSc
- Jeff Ecker, MD

NLP for Risk Prediction in Maternity Care

Mark Clapp, MD MPH
Massachusetts General Hospital
Harvard Medical School

mark.clapp@mgh.harvard.edu @markaclapp

Perelman School of Medicine
University of Pennsylvania

Pregnancy in the Age of the Internet: A Content Analysis of Online Pregnancy Forums

AAMC Maternal Health Equity Workshop | May 18, 2023

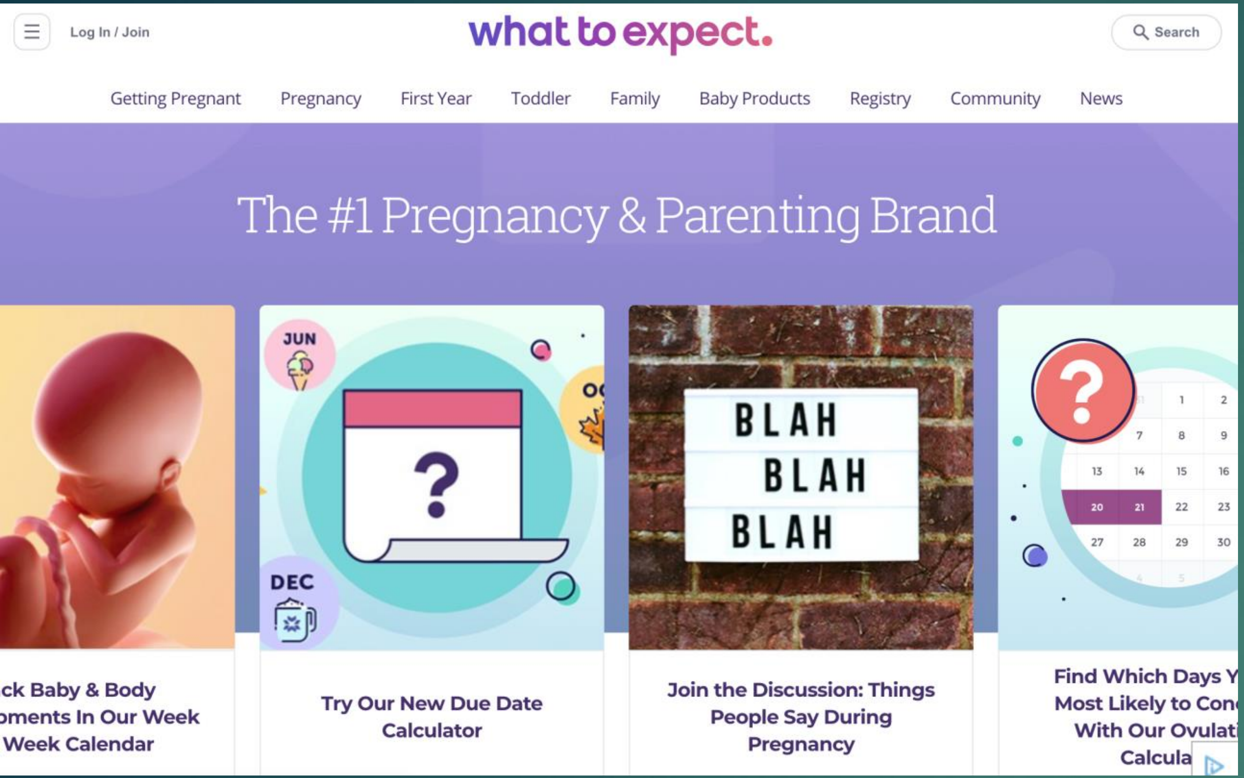
Anna Wexler
Assistant Professor, Department of Medical Ethics & Health Policy
University of Pennsylvania Perelman School of Medicine

PREGNANCY AND DIGITAL HEALTH

- Pregnant individuals increasingly turn to the Internet & digital health during pregnancy (Wexler et al. 2020)

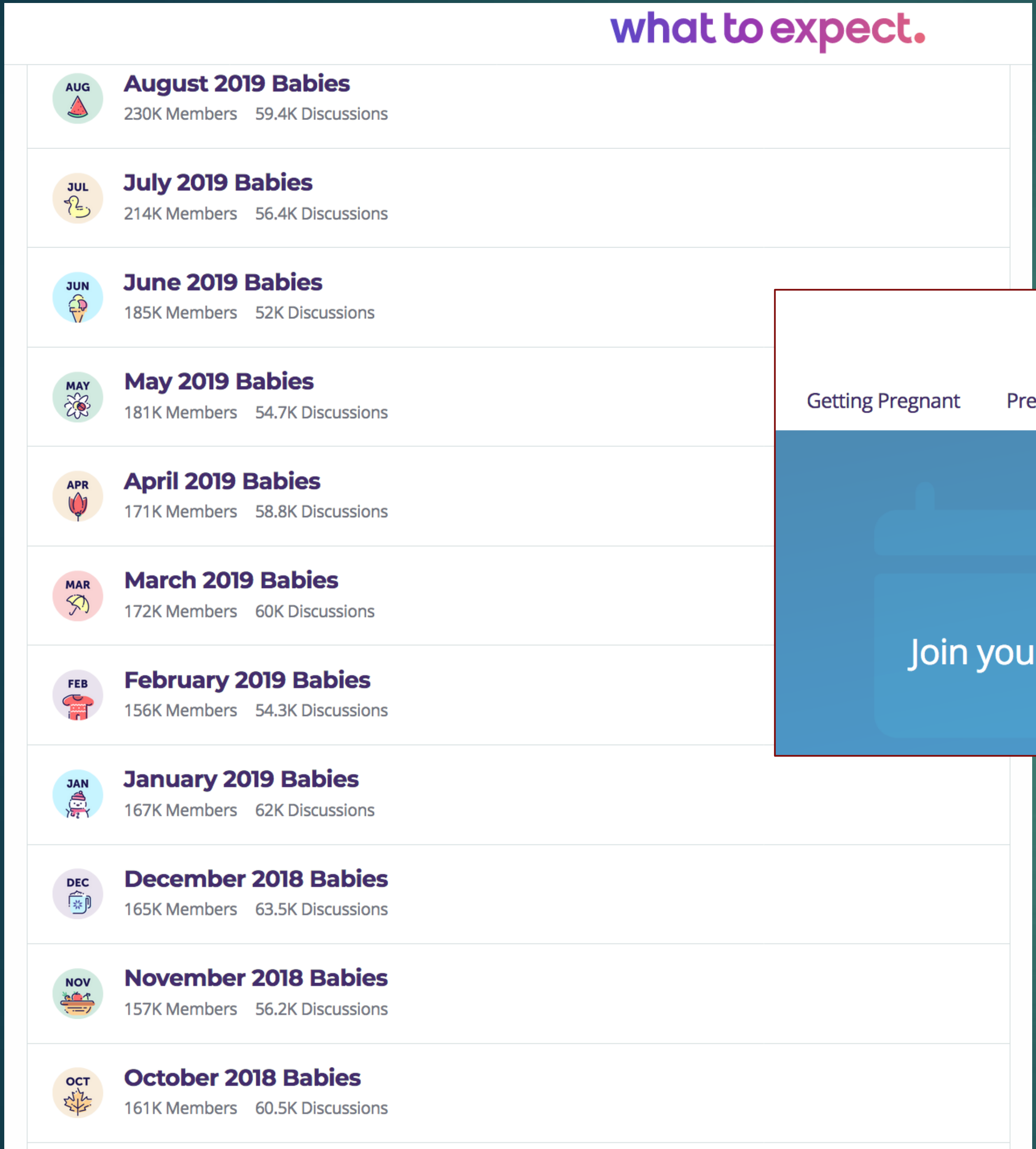
- Prior work has assessed how and why people utilize online health information during pregnancy: they find it empowering, entertaining, quick, reassuring (Sayakhot & Carolan-Olah 2016, for review)

- **Yet little work has assessed what pregnant individuals generate (i.e., online content)**


Log In / Join
what to expect.
Search
Getting Pregnant
Pregnancy
First Year
Toddler
Family
Baby Products
Registry
Community
News
The #1 Pregnancy & Parenting Brand
JUN
DEC
BLAH
BLAH
BLAH
ck Baby & Body
oments In Our Week
Week Calendar
Try Our New Due Date Calculator
Join the Discussion: Things People Say During Pregnancy
Find Which Days Y
Most Likely to Con
With Our Ovulati
Calcula


babycenter
LOG IN / SIGN UP
COMMUNITY
GETTING PREGNANT
PREGNANCY
BABY
TODDLER
PRESCHOOLER
BIG KID
HEALTH
VIDEO
PRODUCTS & GEAR
FOR YOU
Join now to personalize
BabyCenter for your pregnancy
What does my baby look like now?
SHOW ME
Track your baby's development
Get expert guidance from the world's #1 pregnancy and parenting resources, delivered via email, our apps, and website.
TRACK MY BABY


what to expect.
Sign Up
Community Groups
Join our community for baby name ideas, due date discussions, local birth and parent community groups, and more!
Find Moms Who Share Your Birth Month
October 2022
September 2022
August 2022
July 2022
June 2022
May 2022
April 2022
March 2022
February 2022



~500,000 total posts per birth club month across 3 trimesters + post-partum

W

Getting a used breast pump from S-I-L?

25 posts

My sister in law is selling a lot of her baby stuff for cheap and so we are going to get much as we can. However, one of the items is a breast pump. I was all for it due to the cost savings (my insurance won't cover it) but now I'm reading on...

Created by WGriggs17
Last comment from lashawn711 23 minutes ago

A

Flu shot or no?

5 posts

Last time I was pregnant the doctor literally scared me into getting a flu shot... said me and baby could die if I didn't... this time around I don't know what to think or do? Anyone have an opinion on this? Thanks

Created by alisonmariez
Last comment from JRickman 28 minutes ago

It doesn't hurt to be nice

8 posts

Or to just keep scrolling .

Created by dcnewmomma
Last comment from DubMam 29 minutes ago

Heading to the ER and crossing fingers

10 posts

I was having back pain on sunday that was significant enough to make an appt in the middle of the night for Monday with my ob. I generally felt malaise and just rotten. She did a cervix check and urine test stating my urine had some white blood...

Created by ArtisanEGG
Last comment from BB64 31 minutes ago

L

show me your nursery ❤😭

108 posts

23 weeks and its already almost all set up! Am i crazy ?! this is baby #5 and i dont think ive ever got it done this soon ❤ it will be shared with his older sister who will be 19 months when he arrives. i still have alot of organizing to do...

Created by lemonleefresh
Last comment from Lisa2ndtimeprego 31 minutes ago

Jalen or Jaylen —HELP!

33 posts

Help!! I need help deciding on the spelling for my little guy. We like both but can't decide. Thanks!

Created by JMJ1023

Update after a scary situation

55 posts

Hey ! I got discharged from the hospital this afternoon and before I left I spent time with Noah and he's doing good! His breathing tube got moved or clogged whatever happened he stopped breathing so they started pumping air with a bag and his...

Created by MHandy82518
Last comment from Amberlee317 42 minutes ago

Weird gush of fluid

34 posts

So first off i have called the dr. It was the first thing i did.I was at work just standing up and i felt a gush similar to heavy discharge . I went to the bathroom though and it had left quite a patch on my leggings. I will mention I've been...

Created by Hayleyuk
Last comment from DubMam 45 minutes ago

C

Group B strep in urine- concerning?

1 posts

Just got my urine test results with the above diagnosis. Anyone else had such situation?

Created by Catherine123Go 46 minutes ago

Debilitating Round Ligament Pain

1 posts

Anyone else struggling with debilitating round ligament pain? I've been pretty much bed bound for three days now and I can't find any information online or with my doctor for how to cope. Baths, stretches, Tylenol, belly band, etc. all provide...

Created by katiemroo 47 minutes ago

L

HELP! Abdominal muscle pain!

4 posts

I have been experiencing a slowly increasing pain in what feels like my abdominal muscle on the right side. I can push into my belly and find a VERY painful spot and it seems to get worse when I am active or working (working especially, I'm...

Created by Leandra1986
Last comment from katiemroo 48 minutes ago

1

Iron supplements

12 posts

My doctor told me I need to take an iron supplement in addition to my (presciption) prenatal that has extra iron in it.I asked her twice but I've never taken iron before and I'm still

- Goal: to better understand the topics that individuals post about on public online pregnancy forums using automatic methods of language analysis (specifically, topic modeling)
  - to discover the extent to which individuals discuss health-related topics, and which topics appear most frequently
  - to differentially characterize the topics that individuals discuss across the three trimesters and the postpartum period

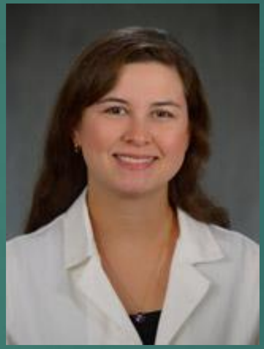
THE TEAM

Graciela Gonzalez-Hernandez
Associate Professor of Informatics
Department of Biostatistics, Epidemiology, and Informatics

Karen O’Connor
Senior Data Analyst
Department of Biostatistics, Epidemiology, and Informatics

Davy Weissenbacher
Research Data Scientist
Department of Biostatistics, Epidemiology, and Informatics

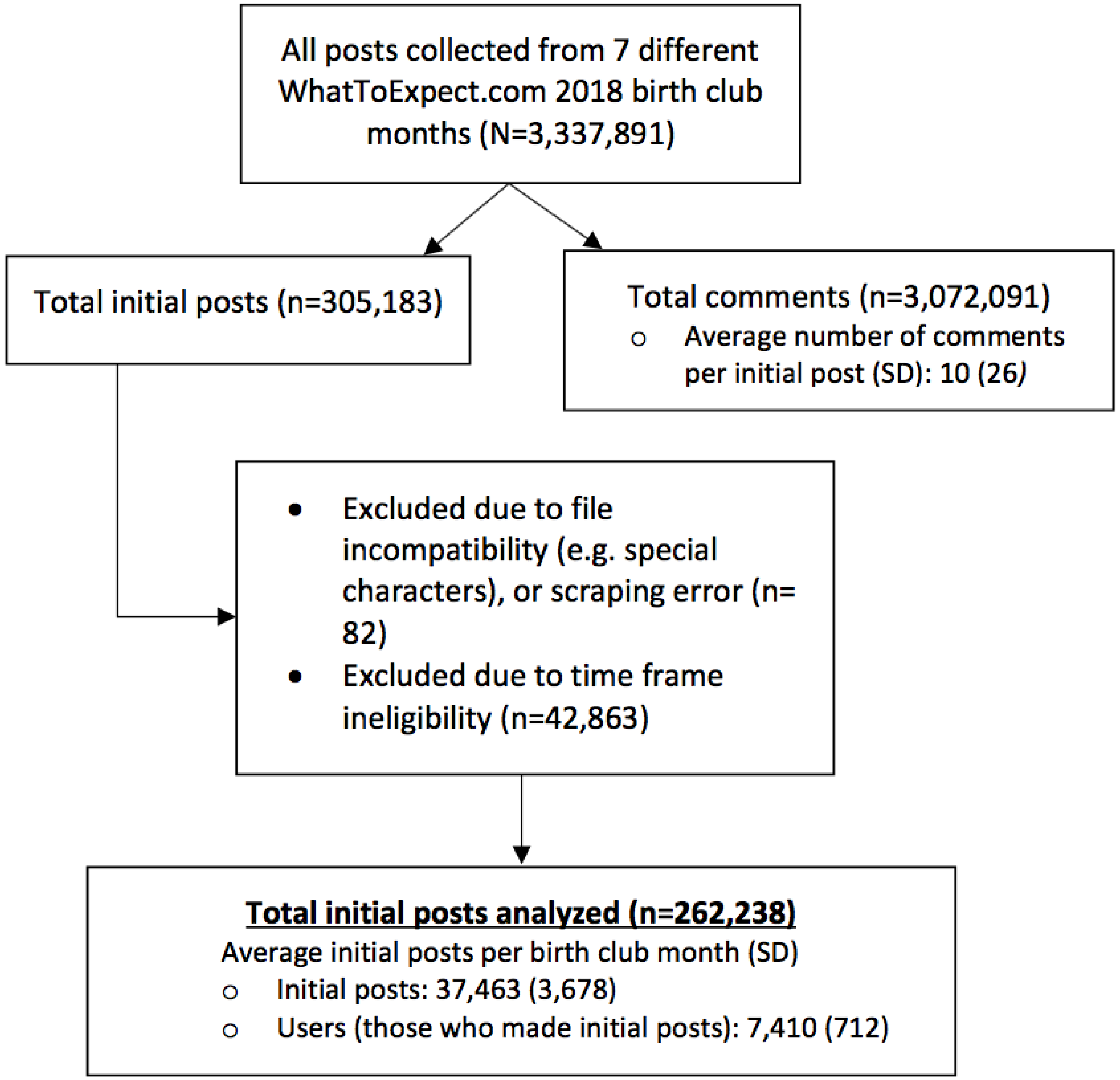
Anahita Davoudi
Lead Statistician
Department of Biostatistics, Epidemiology, and Informatics



Holly Cummings
Assistant Professor of Clinical Obstetrics and Gynecology
Department of Obstetrics and Gynecology

Rebekah Choi
Project Manager
Department of Medical Ethics and Health Policy

Figure 1. **Data collection flowchart.**



TOPIC MODELING: LDA

1. Topic modeling method known as Latent Dirichlet Allocation (LDA), which discovers a “word cluster” that has a high probability of appearing together

Example: '0.068*"morning" + 0.050*"symptom" + 0.040*"nausea" + 0.037*"pregnancy" + 0.035*"throw" + 0.033*"trimester" + 0.032*"stomach" + 0.030*"sickness" + 0.030*"sick" + 0.023*"nauseous" + 0.018*"vomit" + 0.017*"pregnant" + 0.017*"eat" + 0.016*"experience" + 0.014*"wrong"'

2. Requires pre-processing data & setting parameters (50 topics, 200 iterations)

Figure 3. Distribution of posts by label and rank per time period

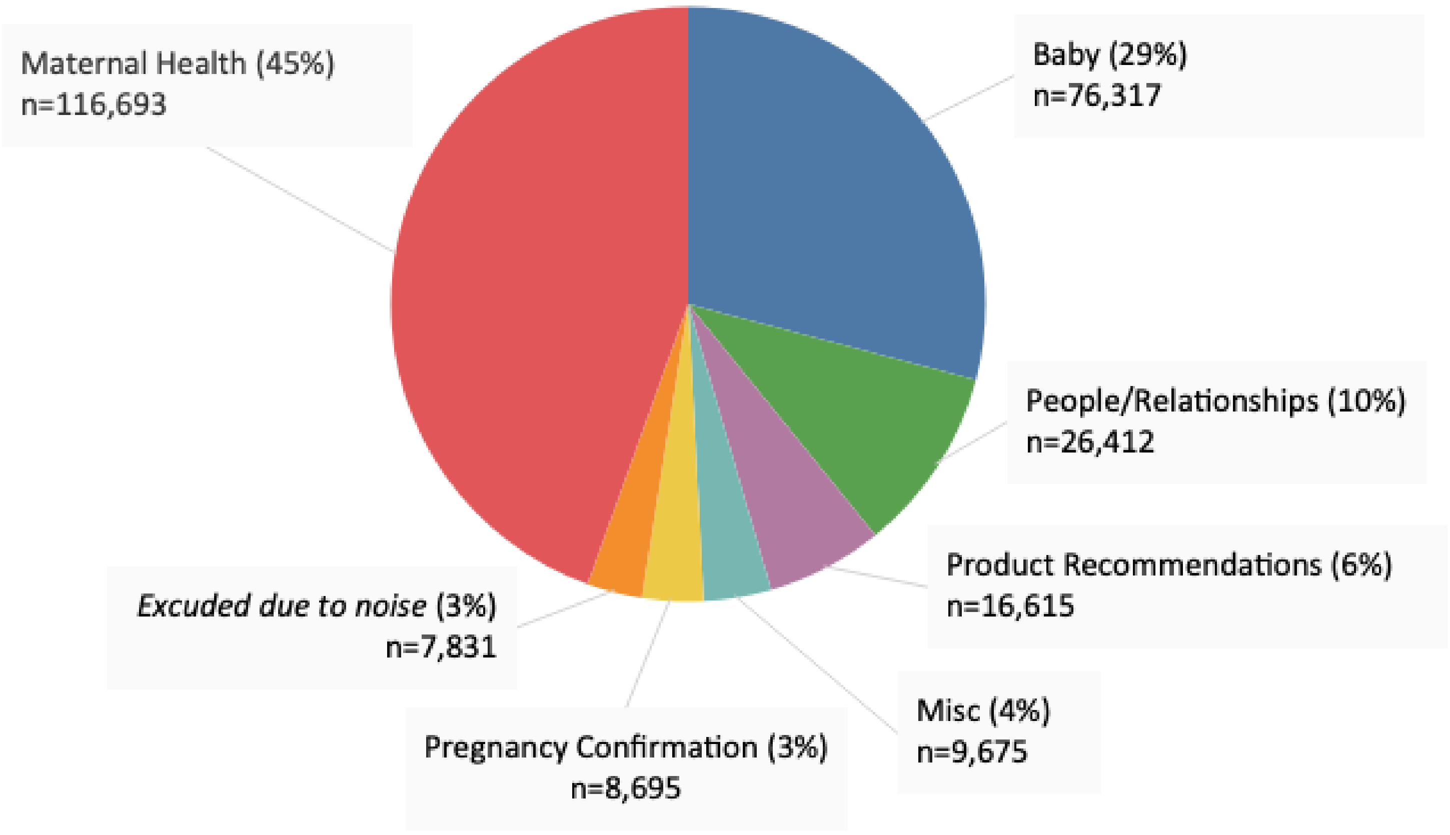
Category	Subcategory	Topic Label	Total Posts	Word Cluster	T1	T2	T3	PP
Maternal Health	Physical Symptoms	Pain (mostly abdomen/back/pelvic)	13,196	pain, back, sit, hurt, experience, walk, pressure, stomach, stand, pelvic, help, hip				
		Morning sickness/nausea	9,751	morning, symptom, nausea, pregnancy, throw, trimester, stomach, sickness, sick, nauseous, vomit, pregnant				
		Bleeding/spotting/cramping	9,357	cramp, bleed, period, spot, blood, discharge, normal, light, experience, red, brown, wipe				
		Misc health symptoms (e.g., sore/painful breasts, constipation, hemorrhoids)	5,488	weird, pee, hurt, sore, boob, happen, leak, normal, burn, tmi, heartburn, notice				
		Anxiety/depression; medication questions	4,798	anxiety, pregnancy, help, postpartum, tear, pregnant, infection, issue, doctor, experience, effect, medication				
		Cold/flu symptoms	4,661	sick, cold, cough, flu, fever, shoot, nose, breathe, throat, catch, ear, help				
	Labor	Labor (induction/dilation)	12,062	induce, cm, dilate, labor, push, check, induction, cervix, due, epidural, doctor, bear				
		Labor signs (contractions, am I in labor)	6,629	contraction, labor, water, break, min, braxton, apart, sign, hick, painful, hospital, strong				
		Anxiety about delivery	5,930	birth, hospital, nervous, scare, delivery, deliver, plan, experience, labor, natural, mom, prepare				
		Mucus plug (& baby position/breech q's)	1,989	head, show, lose, turn, mucus, plug, bloody, position, breech, flip, flat, mucous				
	Body and mood changes	Diet & food (cravings, aversions, safety)	6,578	eat, drink, food, water, meal, dinner, tea, crave, snack, coffee, lunch, ice				
		Stretch marks & baby bump	5,558	belly, bump, leg, stretch, mark, pop, button, itchy, oil, pregnancy, bloat, itch				
		Weight (gain/loss)	4,584	weight, gain, lb, pound, lose, pregnancy, weigh, eat, healthy, pregnant, pre, fat				
		Energy levels (fatigure) & exercise	3,962	pregnancy, body, run, tire, pregnant, walk, struggle, toddler, energy, exhaust, remember, easy				
		Sex (sex drive, sex after birth)	1,913	crazy, sex, drive, headache, hudband, pregnant, breath, hubby, migraine, pregnancy, nut, insane				
	Pregnancy complications	Pre-eclampsia (BP/protein/urine) but also general doc appt	9,209	doctor, appointment, blood, ob, dr, check, pressure, office, appt, back, nurse, concern				
		Placenta disorders during pregnancy (& hypnobirthing/natural birth q's)	2,132	hear, curious, experience, opinion, placenta, wonder, vs, whether, mum, class, love, mom				
Baby	Preparing for baby	Excited about baby	8,981	girl, boy, excite, love, sweet, meet, beautiful, amaze, healthy, son, team, bless				
		Choosing baby name	8,319	name, middle, girl, boy, love, husband, pick, help, decide, suggestion, choose, opinion				
		First ultrasound picture & early gender prediction	5,343	picture, pic, foot, swell, ramzi, twin, dream, photo, ultrasound, theory, nub, skull				
		Gender reveal	3,171	guess, gender, fun, reveal, csection, surprise, ultrasound, wife, scan, idea, tale, correct				
	Baby's health (in utero)	Concerns about miscarriage	8,896	pregnancy, miscarriage, heartbeat, ultrasound, level, pregnant, hcg, sac, nervous, progesterone, hear, pray				
		Concerns about ultrasound results	8,414	ultrasound, scan, measure, anatomy, doctor, tech, small, growth, experience, result, back, fluid				
		Baby movement	4,625	move, kick, movement, tummy, roll, active, count, back, crawl, normal, lay, doe				
		Baby heart rate	833	heart, rate, beat, glass, cup, wine, program, sippy, babys, proud, bpm, star				
	Caring for newborn	Baby sleep routine	10,573	sleep, night, wake, nap, fall, asleep, back, bed, awake, help, eat, fee				
		Pumping & breastfeeding	9,305	pump, breastfeed, milk, breast, feed, nurse, bottle, supply, nipple, fee, formula, latch				
		Baby feeding routines	5,896	bottle, formula, oz, eat, feed, spit, gas, solid, reflux, fee, burp, switch				
		Baby sleep arrangements	4,505	room, crib, play, sleep, swaddle, bassinet, rock, nursery, transition, arm, bed, pack				
		Skin problems (mostly baby, some mom)	3,443	change, diaper, face, skin, rash, app, iron, help, wet, acne, red, cream				
		Baby teething & thumb sucking	1,275	hand, mouth, suck, tooth, book, teethe, finger, pull, toy, pacifier, cut, lip				
		Bowel movements (some pet related)	976	poop, dog, video, cat, pooping, potty, green, train, pooped, ride, fresh, normal				
	Misc	Baby eye color (also pink eye; nursery room color)	658	eye, open, white, color, blue, black, dark, pink, light, wide, fill, contact				
People/relationships		Husband (related to household chores, work, kids)	13,478	husband, house, kid, work, help, back, care, alone, son, stress, vent, mom				
		Family/friends	12,934	family, friend, mom, child, people, husband, pregnant, mother, sister, parent, law, daughter				
Product recommendations		Buying maternity clothes & baby items	7,325	buy, maternity, bag, size, clothe, stuff, diaper, brand, store, pack, item, newborn				
		Baby transport options (carseats, strollers)	6,343	car, seat, job, pay, stroller, insurance, cover, school, money, double, infant, plan				
		Bra recommendations	2,947	wear, clean, wash, fit, bra, dress, bath, clothe, nurse, pad, wrap, free				
Pregnancy confirmation		Pregnancy test (also glucose test)	4,643	test, positive, line, result, negative, blood, pregnancy, glucose, faint, sugar, morning, fail				
		Estimating due date & trying to conceive	4,052	due, june, april, september, march, pregnant, january, cycle, bfp, late, august, july				
Misc		Emotional moms & crying/fussy babies	4,225	cry, hold, scream, hate, stop, calm, help, smile, love, sad, leap, upset				
		Travel advice	2,758	advice, tip, appreciate, plan, live, travel, trip, help, weekened, fly, husband, vacation				
		Misc temperature (hot weather, hot flashes, dressing baby for weather)	1,430	safe, hot, cold, warm, heat, cool, summer, outside, sweat, air, winter, weather				
		Back to work & childcare plans	1,262	work, back, schedule, full, daycare, shift, routine, mom, growth, spurt, hr, overnight				
		Six topics excluded due to noise	7,831					
		TOTAL	262,238					

T1 = First trimester
T2 = Second trimester
T3 = Third trimester
PP = Postpartum period

Rank Per Time Period
44.00 1.00

Post Counts
658 13,478

Figure 2. Distribution of posts by general category



Category	Subcategory	Topic Label	Total Posts	Word Cluster	T1	T2	T3	PP
Maternal Health	Physical Symptoms	Pain (mostly abdomen/back/pelvic)	13,196	pain, back, sit, hurt, experience, walk, pressure, stomach, stand, pelvic, help, hip				
		Morning sickness/nausea	9,751	morning, symptom, nausea, pregnancy, throw, trimester, stomach, sickness, sick, nauseous, vomit, pregnant				
		Bleeding/spotting/cramping	9,357	cramp, bleed, period, spot, blood, discharge, normal, light, experience, red, brown, wipe				
		Misc health symptoms (e.g., sore/painful breasts, constipation, hemorrhoids)	5,488	weird, pee, hurt, sore, boob, happen, leak, normal, burn, tmi, heartburn, notice				
		Anxiety/depression; medication questions	4,798	anxiety, pregnancy, help, postpartum, tear, pregnant, infection, issue, doctor, experience, effect, medication				
		Cold/flu symptoms	4,661	sick, cold, cough, flu, fever, shoot, nose, breathe, throat, catch, ear, help				
	Labor	Labor (induction/dilation)	12,062	induce, cm, dilate, labor, push, check, induction, cervix, due, epidural, doctor, bear				
		Labor signs (contractions, am I in labor)	6,629	contraction, labor, water, break, min, braxton, apart, sign, hick, painful, hospital, strong				
		Anxiety about delivery	5,930	birth, hospital, nervous, scare, delivery, deliver, plan, experience, labor, natural, mom, prepare				
		Mucus plug (& baby position/breech q's)	1,989	head, show, lose, turn, mucus, plug, bloody, position, breech, flip, flat, mucous				
	Body and mood changes	Diet & food (cravings, aversions, safety)	6,578	eat, drink, food, water, meal, dinner, tea, crave, snack, coffee, lunch, ice				
		Stretch marks & baby bump	5,558	belly, bump, leg, stretch, mark, pop, button, itchy, oil, pregnancy, bloat, itch				
		Weight (gain/loss)	4,584	weight, gain, lb, pound, lose, pregnancy, weigh, eat, healthy, pregnant, pre, fat				
		Energy levels (fatigure) & exercise	3,962	pregnancy, body, run, tire, pregnant, walk, struggle, toddler, energy, exhaust, remember, easy				
		Sex (sex drive, sex after birth)	1,913	crazy, sex, drive, headache, hudband, pregnant, breath, hubby, migraine, pregnancy, nut, insane				
	Pregnancy complications	Pre-eclampsia (BP/protein/urine) but also general doc appt	9,209	doctor, appointment, blood, ob, dr, check, pressure, office, appt, back, nurse, concern				
		Placenta disorders during pregnancy (& hypnobirthing/natural birth q's)	2,132	hear, curious, experience, opinion, placenta, wonder, vs, whether, mum, class, love, mom				

T1 = First trimester
T2 = Second trimester
T3 = Third trimester
PP = Postpartum period

Rank Per Time Period
44.00 1.00

Post Counts
658 13,478

Baby	Preparing for baby	Excited about baby	8,981	girl, boy, excite, love, sweet, meet, beautiful, amaze, healthy, son, team, bless
		Choosing baby name	8,319	name, middle, girl, boy, love, husband, pick, help, decide, suggestion, choose, opinion
		First ultrasound picture & early gender prediction	5,343	picture, pic, foot, swell, ramzi, twin, dream, photo, ultrasound, theory, nub, skull
		Gender reveal	3,171	guess, gender, fun, reveal, csection, surprise, ultrasound, wife, scan, idea, tale, correct
	Baby's health (in utero)	Concerns about miscarriage	8,896	pregnancy, miscarriage, heartbeat, ultrasound, level, pregnant, hcg, sac, nervous, progesterone, hear, pray
		Concerns about ultrasound results	8,414	ultrasound, scan, measure, anatomy, doctor, tech, small, growth, experience, result, back, fluid
		Baby movement	4,625	move, kick, movement, tummy, roll, active, count, back, crawl, normal, lay, doe
		Baby heart rate	833	heart, rate, beat, glass, cup, wine, program, sippy, babys, proud, bpm, star
	Caring for newborn	Baby sleep routine	10,573	sleep, night, wake, nap, fall, asleep, back, bed, awake, help, eat, fee
		Pumping & breastfeeding	9,305	pump, breastfeed, milk, breast, feed, nurse, bottle, supply, nipple, fee, formula, latch
		Baby feeding routines	5,896	bottle, formula, oz, eat, feed, spit, gas, solid, reflux, fee, burp, switch
		Baby sleep arrangements	4,505	room, crib, play, sleep, swaddle, bassinet, rock, nursery, transition, arm, bed, pack
		Skin problems (mostly baby, some mom)	3,443	change, diaper, face, skin, rash, app, iron, help, wet, acne, red, cream
		Baby teething & thumb sucking	1,275	hand, mouth, suck, tooth, book, teethe, finger, pull, toy, pacifier, cut, lip
		Bowel movements (some pet related)	976	poop, dog, video, cat, pooping, potty, green, train, pooped, ride, fresh, normal
	Misc	Baby eye color (also pink eye; nursery room color)	658	eye, open, white, color, blue, black, dark, pink, light, wide, fill, contact

T1 = First trimester
T2 = Second trimester
T3 = Third trimester
PP = Postpartum period

Rank Per Time Period
44.00 1.00

Post Counts
658 13,478

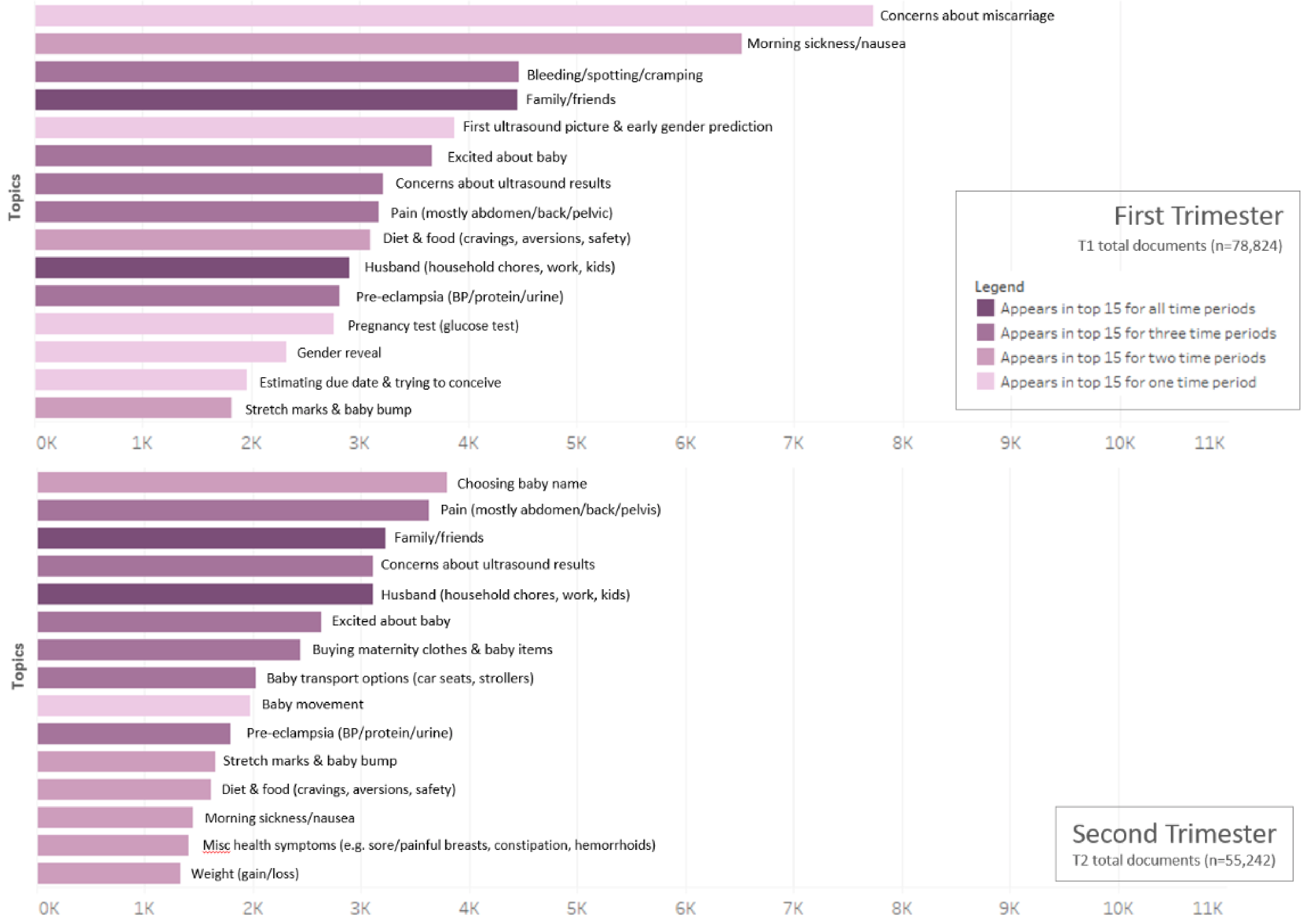
People/relationships	Husband (related to household chores, work, kids)	13,478	husband, house, kid, work, help, back, care, alone, son, stress, vent, mom
	Family/friends	12,934	family, friend, mom, child, people, husband, pregnant, mother, sister, parent, law, daughter
Product recommendations	Buying maternity clothes & baby items	7,325	buy, maternity, bag, size, clothe, stuff, diaper, brand, store, pack, item, newborn
	Baby transport options (carseats, strollers)	6,343	car, seat, job, pay, stroller, insurance, cover, school, money, double, infant, plan
	Bra recommendations	2,947	wear, clean, wash, fit, bra, dress, bath, clothe, nurse, pad, wrap, free
Pregnancy confirmation	Pregnancy test (also glucose test)	4,643	test, positive, line, result, negative, blood, pregnancy, glucose, faint, sugar, morning, fail
	Estimating due date & trying to conceive	4,052	due, june, april, september, march, pregnant, january, cycle, bfp, late, august, july
Misc	Emotional moms & crying/fussy babies	4,225	cry, hold, scream, hate, stop, calm, help, smile, love, sad, leap, upset
	Travel advice	2,758	advice, tip, appreciate, plan, live, travel, trip, help, weekend, fly, husband, vacation
	Misc temperature (hot weather, hot flashes, dressing baby for weather)	1,430	safe, hot, cold, warm, heat, cool, summer, outside, sweat, air, winter, weather
	Back to work & childcare plans	1,262	work, back, schedule, full, daycare, shift, routine, mom, growth, spurt, hr, overnight
	Six topics excluded due to noise	7,831	
	TOTAL	262,238	

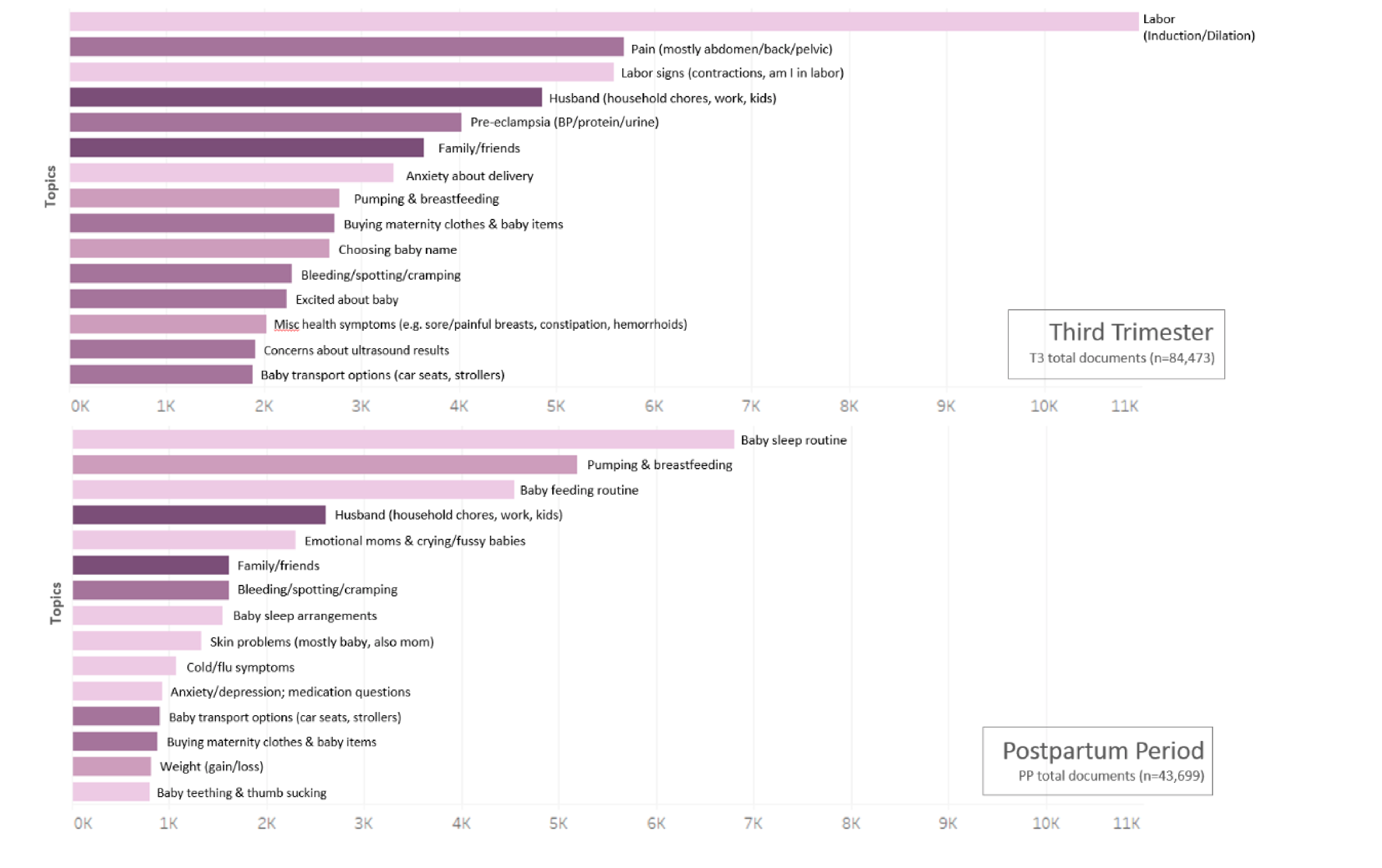
T1 = First trimester
T2 = Second trimester
T3 = Third trimester
PP = Postpartum period

Rank Per Time Period
44.00 1.00

Post Counts
658 13,478

Figure 4. Top 15 topics per time period





Topics
Labor (Induction/Dilation)
Pain (mostly abdomen/back/pelvic)
Labor signs (contractions, am I in labor)
Husband (household chores, work, kids)
Pre-eclampsia (BP/protein/urine)
Family/friends
Anxiety about delivery
Pumping & breastfeeding
Buying maternity clothes & baby items
Choosing baby name
Bleeding/spotting/cramping
Excited about baby
Misc health symptoms (e.g. sore/painful breasts, constipation, hemorrhoids)
Concerns about ultrasound results
Baby transport options (car seats, strollers)
Third Trimester
T3 total documents (n=84,473)
0K 1K 2K 3K 4K 5K 6K 7K 8K 9K 10K 11K
Topics
Baby sleep routine
Pumping & breastfeeding
Baby feeding routine
Husband (household chores, work, kids)
Emotional moms & crying/fussy babies
Family/friends
Bleeding/spotting/cramping
Baby sleep arrangements
Skin problems (mostly baby, also mom)
Cold/flu symptoms
Anxiety/depression; medication questions
Baby transport options (car seats, strollers)
Buying maternity clothes & baby items
Weight (gain/loss)
Baby teething & thumb sucking
Postpartum Period
PP total documents (n=43,699)
0K 1K 2K 3K 4K 5K 6K 7K 8K 9K 10K 11K

DISCUSSION

- More than just emotional support or product recommendations, pregnant individuals are talking extensively about their health on online pregnancy forums

- Areas of greatest health information-seeking

  - Miscarriage
  - Labor
  - Newborn care

- Pregnant individuals report that online information affects their health-care decision-making (Lagan 2010)… yet as many as 70-75% of them do not speak to health care providers about information retrieved from the Internet (Fredriksen et al 2016; Larsson 2009; Gao et al. 2013)

LIMITATIONS

- No demographic information about who participates in these forums
- Topic modeling limitations (i.e., posts that were very general likely did not arise in topic list)

FUTURE QUESTIONS

- Who participates in online pregnancy forums and who does not?
- What are pregnant individuals discussing online that they are not bringing up in their OB's office?
- How is online health information changing the traditional patient-physician relationship in the maternal health context?

Perelman School of Medicine
UNIVERSITY of PENNSYLVANIA

Thank you!

Anna Wexler
awex@pennmedicine.upenn.edu

This project was supported by the Office of the Director, NIH, under Award Number DP5OD026420

Harnessing NLP for Assisting Tele-Triage of Expectant & New Mothers in Kenya

Amulya Yadav
Penn State University


OLD MAIN

1 1 The State of Maternal Health Care in Kenya

<u>**Overview**</u>**:**

- Kenya has some of the highest rates of maternal and newborn mortality
    - 14 times higher than the US
    - 100 times higher than some countries in Western Europe
    - below the international minimum to deliver essential health services

- Why is this happening?
    - Mothers live in rural areas quite far away from hospitals or clinic
    - They lack access to timely information about their medical symptoms

1 1

The State of Maternal Health Care in Kenya

Overview:

- Kenya has some of the highest rates of maternal and newborn mortality
    - 14 times higher than the US
    - 100 times higher than some countries in Western Europe
    - below the international minimum to deliver essential health services

- Why is this happening?
    - Mothers live in rural areas quite far away from hospitals or clinic
    - They lack access to timely information about their medical symptoms
        - Do my current medical symptoms warrant a trip to the doctor?

PennState

Enter Jacaranda Health

- Kenya-based non-profit organization started in 2017.
- They deliver low-cost and sustainable solutions to improve the quality of care for mothers and newborns.
- Currently serving more than 3 million mothers and babies annually.

1 3 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

Users register for PROMPTS

1 3 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

Users register for PROMPTS → Users ask questions about pregnancy

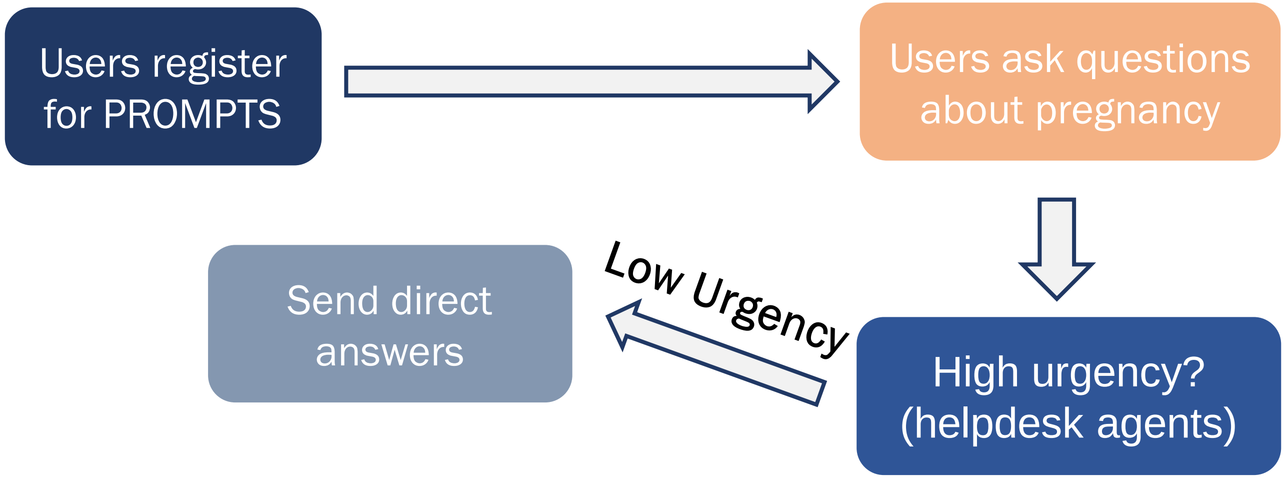
1 3 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

Users register for PROMPTS → Users ask questions about pregnancy → High urgency? (helpdesk agents)

1 3 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.



1 3 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

Users register for PROMPTS

Users ask questions about pregnancy

High urgency? (helpdesk agents)

Low Urgency

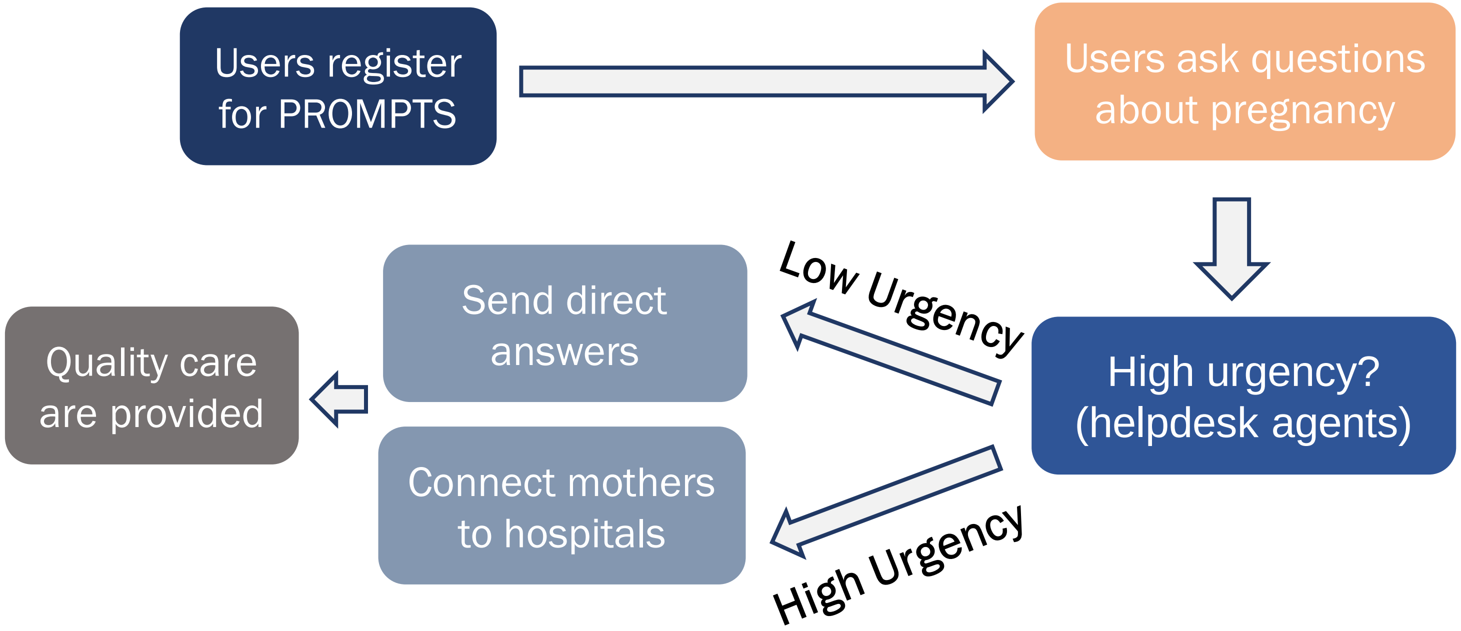
Send direct answers

High Urgency

Connect mothers to hospitals

PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.



13 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

2 million enrolled into PROMPTS so far

0.35 million users enroll to PROMPTS every month

1.1 million messages received every month

13

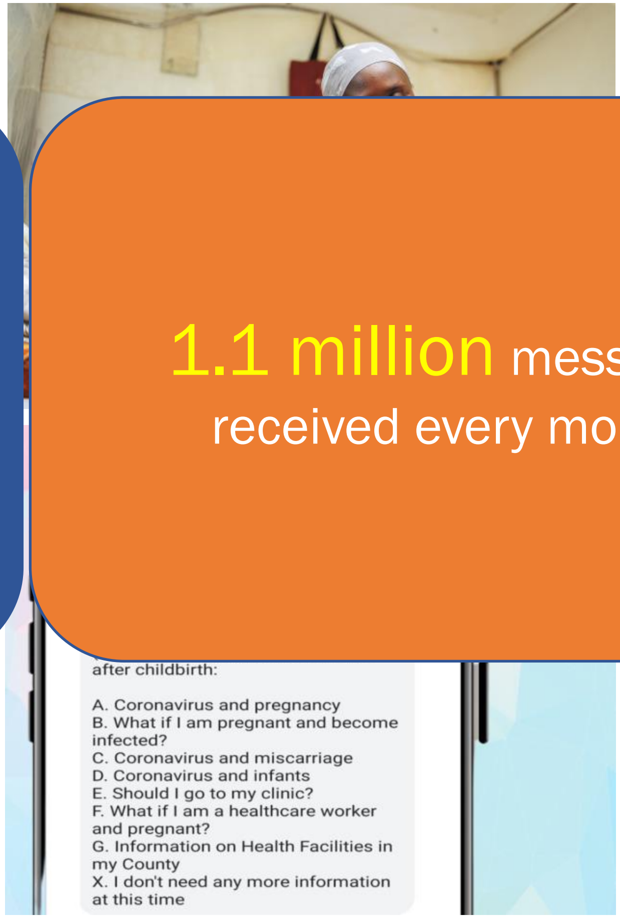
PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

2 million enrolled into PROMPTS so far

0.35 million users enroll to PROMPTS every month

1.1 million messages received every month



1 3 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

2 million enrolled into PROMPTS so far

0.35 million users enroll to PROMPTS every month

1.1 million messages received every month

Answering messages is not straightforward for helpdesk agents

PennState

1 3 PROMPTS

A digital health platform that connects mothers with lifesaving advice and referral to care.

2 million e
PROMPTS

How can we use NLP to improve the health care situation in Kenya and provide a better service experience?

1 million messages
received every month

Answering messages is not straightforward for helpdesk agents

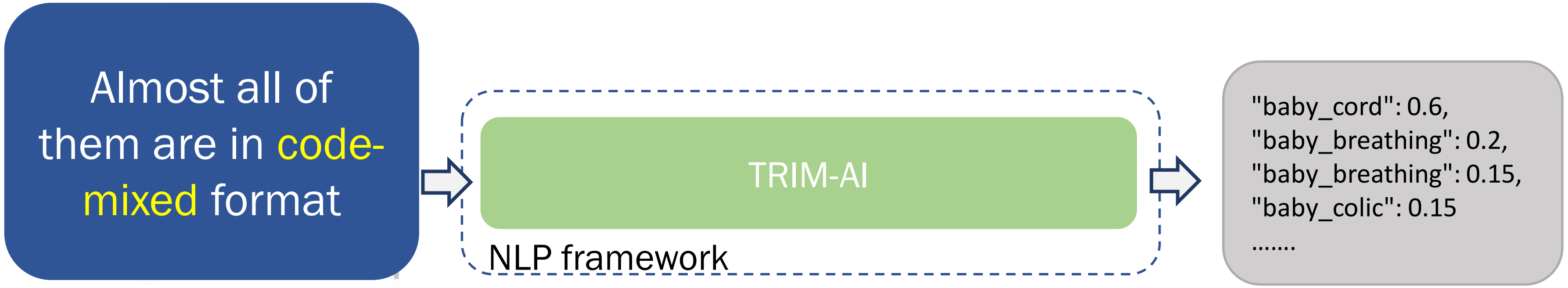
2 1 Problem Statement

Goal:

- Automate the task of predicting emergency levels of a user’s medical condition based on their SMS messages

Contribution:

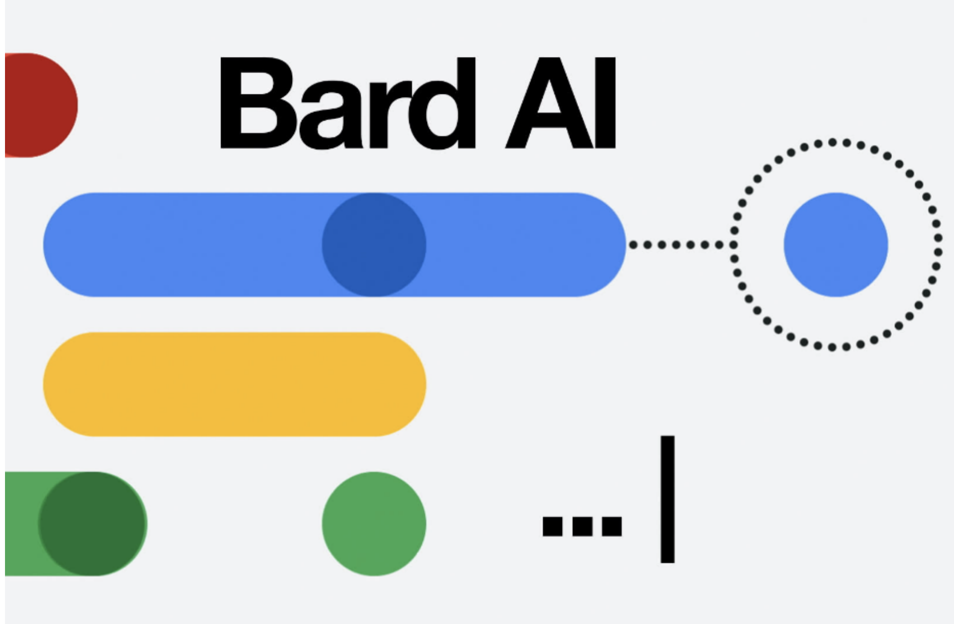
- Propose an NLP framework, TRIM-AI, to classify cases into different emergency levels
- Utilize a special Large Language Model (LLM) and continue pretraining to deal with the code-mixed text
- Output a sorted batch of incoming SMS messages (in decreasing order of risk score)



Input of SMS message

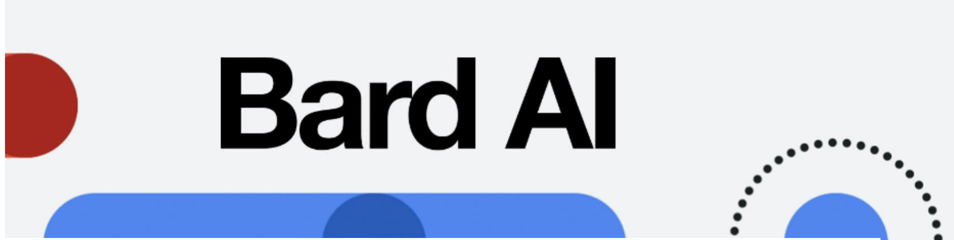
“Intent” Probability Distribution

A Quick Intro on LLMs



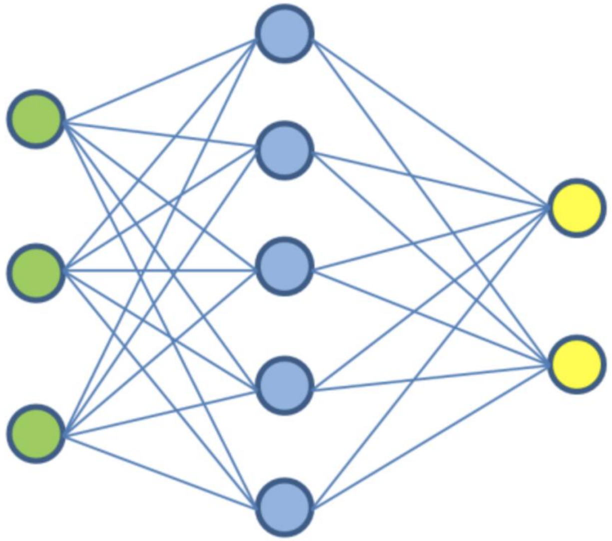
2 3

A Quick Intro on LLMs: What do they do?



A Quick Intro on LLMs: Building Classifiers

Meta LLAMA Model
(An Example of an LLM)



Classification
Network

Predictions

A Quick Intro on LLMs: Building Classifiers

Can we use an LLM based classifier to predict the severity of a mother’s medical condition using the content of their SMS messages?

Predictions

Me

(An Example of an LLM)

Network

2 6 Code-Mixing Problem with LLMs

Code-mixing phenomenon:

Phrases or words which belong to different languages appear in one sentence.

- Blue: English
- Red: Swahili

Example 1:

Code-mixed message: **My son alimeza a coin since sunday na ajaenda haja is it risk ama itatoka tu**

Translated English message: **My son has swallowed a coin since sunday and that is going to need is it risky or it will just go away**

Example 2:

Code-mixed message: **Je heartburn inasababishwa na nini? sababu hata nikule nini lazma nikuwe nayo**

Translated English message: **What causes heartburn? Because even if I eat what I have to have**

27 Code-Mixing Problem with LLMs

Most SMS messages received by the PROMPTS platform are code-mixed between English & Swahili

mple 1:
-mixed messag
da haja is it ris
lated English
y and that is g

ple 2:
-mixed messag
u hata nikule
slated English
because even if I eat w

How do we use LLMs with code-mixed data?

28 Dataset Summary

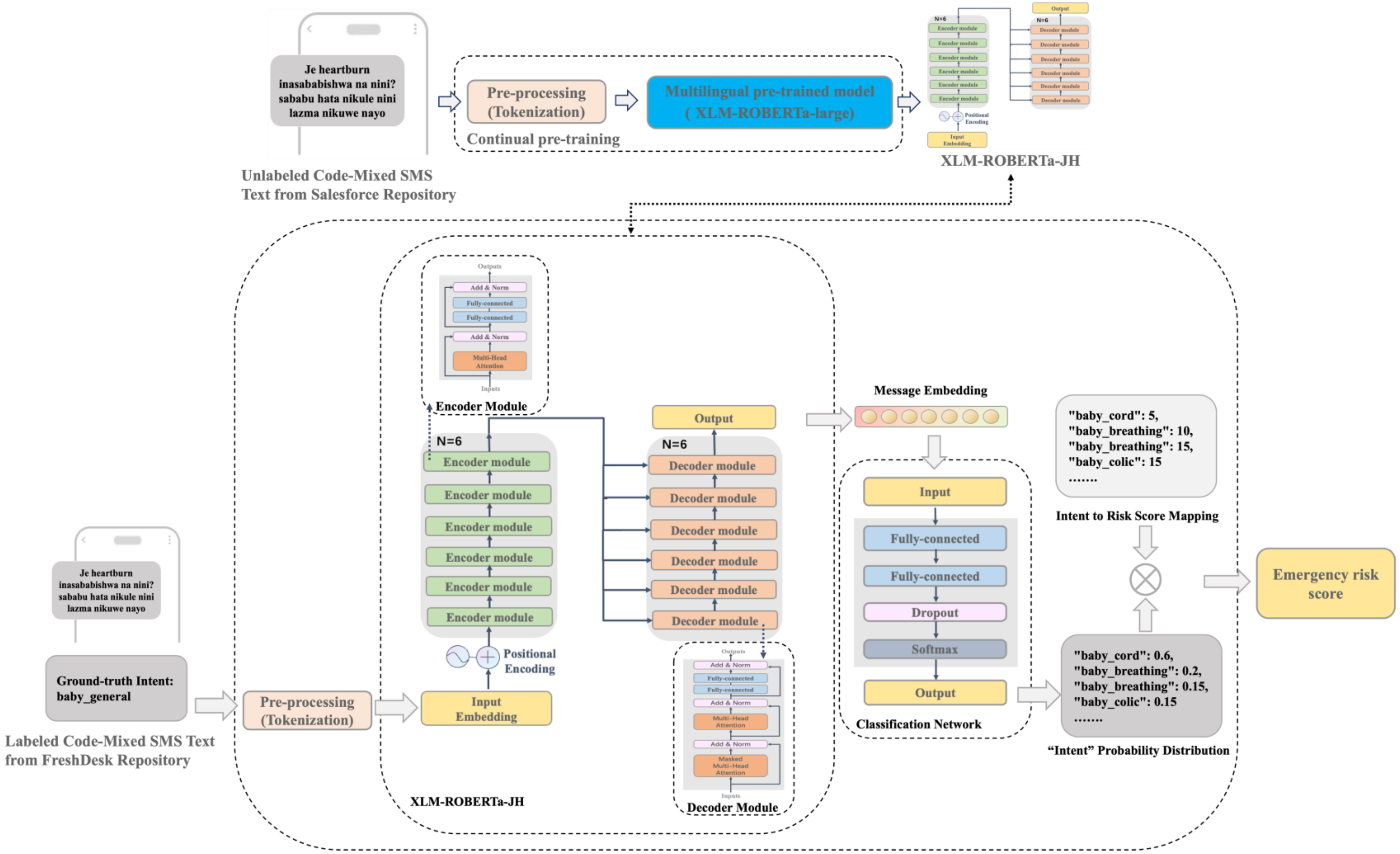
Dataset Statistic:

Code-mixed SMS messages within PROMPTS are stored across two different repositories:

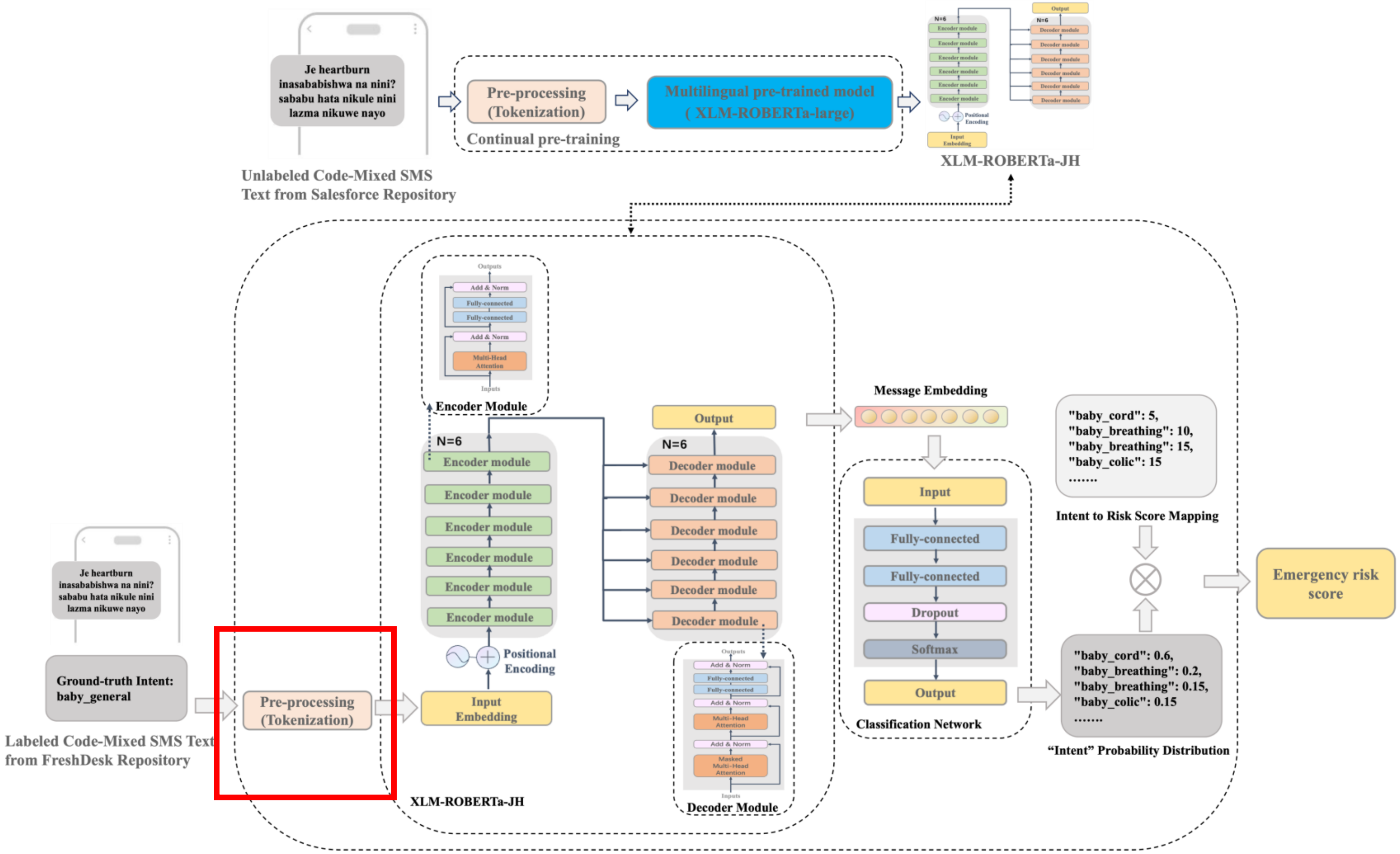
1. SalesForce dataset (Unlabeled)
   - The total number of text messages is: 939,819

2. Freshdesk dataset (Labeled with the correct emergency level)
   - The total number of text messages is: 107,717
   - The total number of intent (emergency level) type is: 58

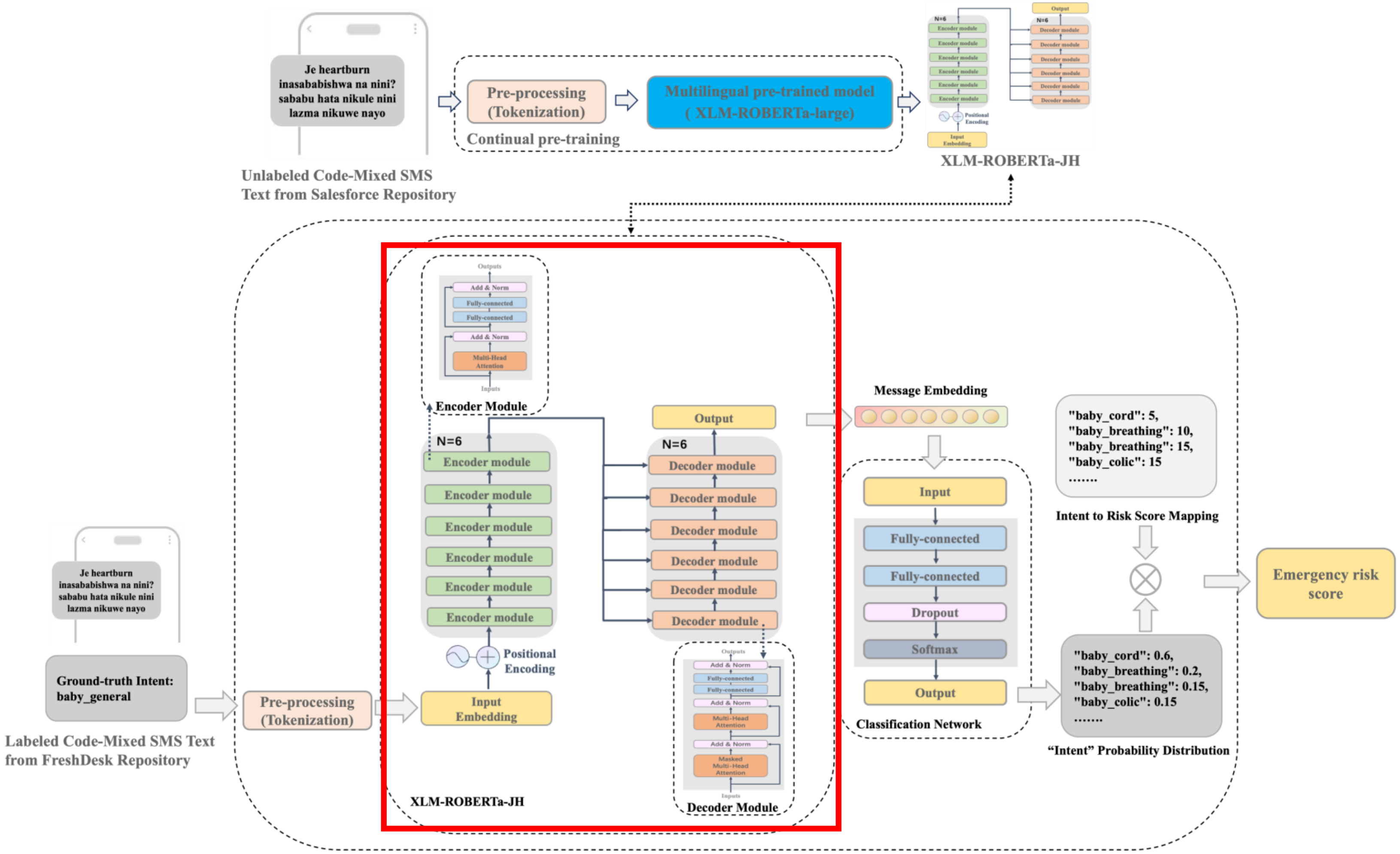
TRIM-AI: Model Architecture



3 2 TRIM-AI: Model Architecture (preprocessing module)

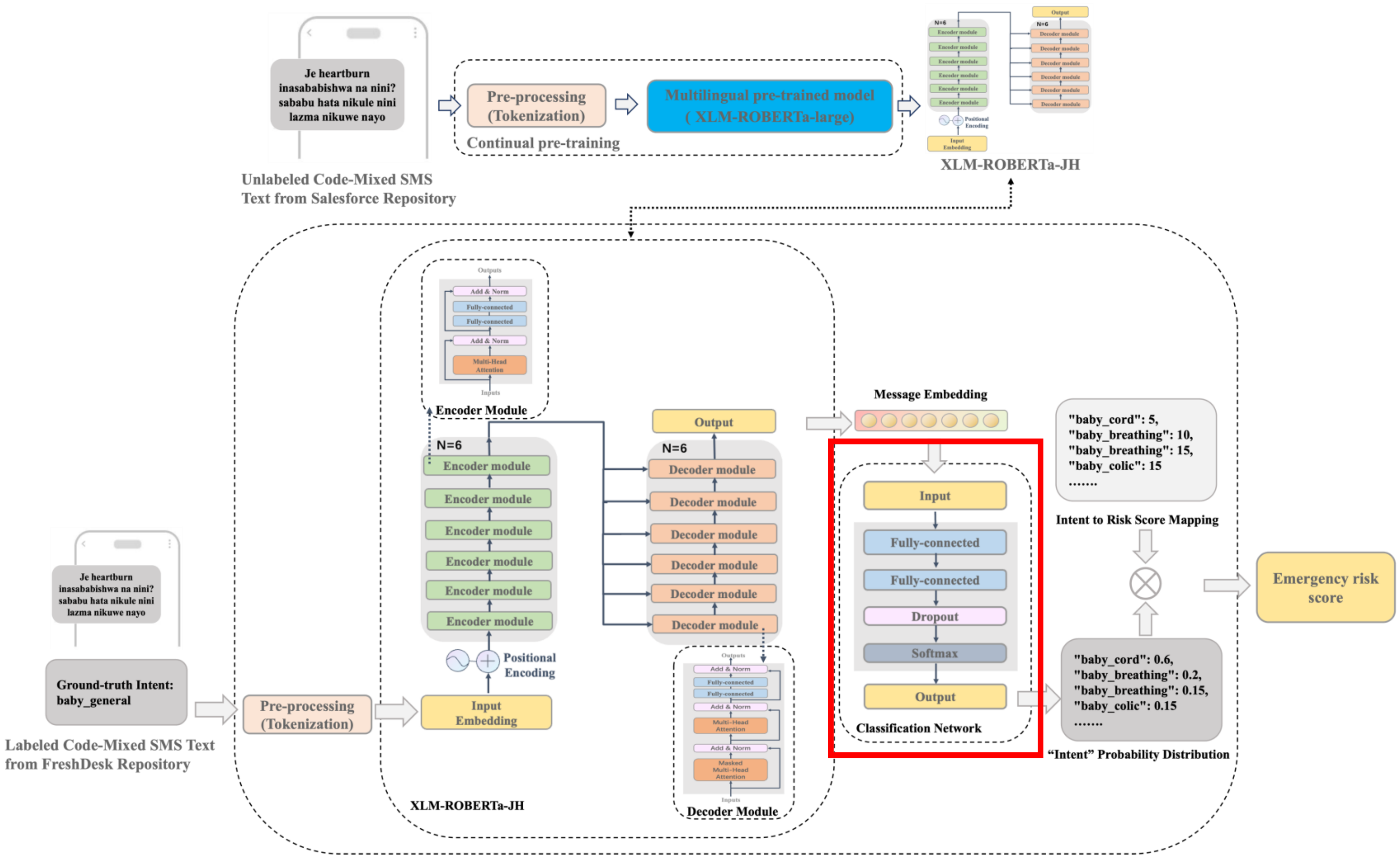


3 3 TRIM-AI: Model Architecture (XLM-ROBERTa-JH)



PennState

3 4 TRIM-AI: Model Architecture (classification network)



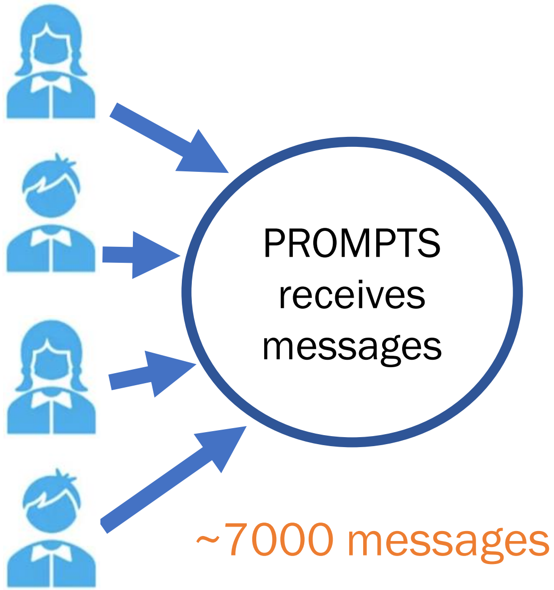
Performance Evaluation

Baseline comparison:

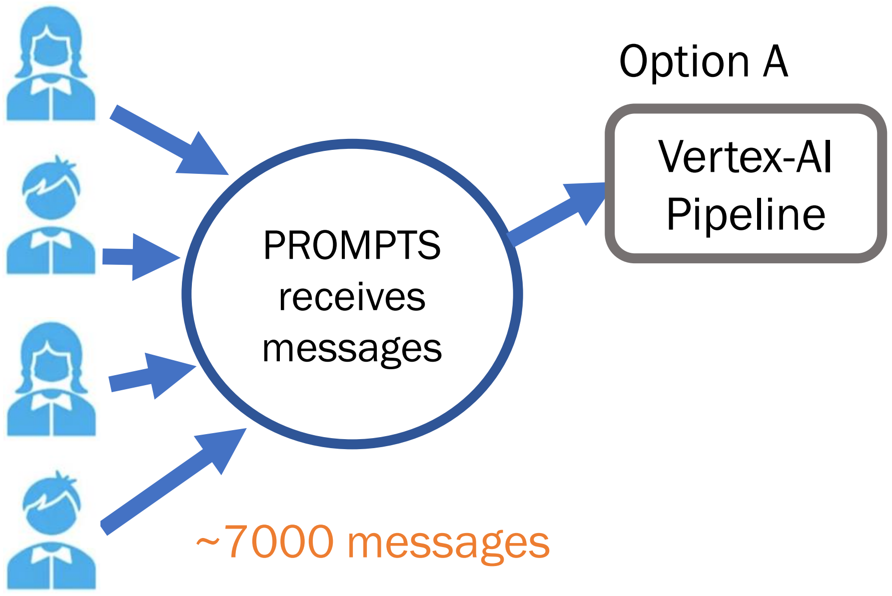
Model Name	Weighted Precision	Weighted Recall	Weighted F_1-score
Hierarchical NN (FastText as the encoder layer)	0.678	0.668	0.669
Hierarchical NN (TextRNN as the encoder layer)	0.682	0.673	0.67
Hierarchical NN (TextCNN as the encoder layer)	0.638	0.629	0.626
Hierarchical NN (RCNN as the encoder layer)	0.679	0.667	0.663
TRIM-AI (monolingual ROBERTa-base)	0.730	0.729	0.728
TRIM-AI (m-BERT)	0.727	0.726	0.725
TRIM-AI (XLM-ROBERTa-base)	0.736	0.735	0.735
Vertex AI model	0.765	0.598	0.671
TRIM-AI	**0.775**	**0.775**	**0.774**

42

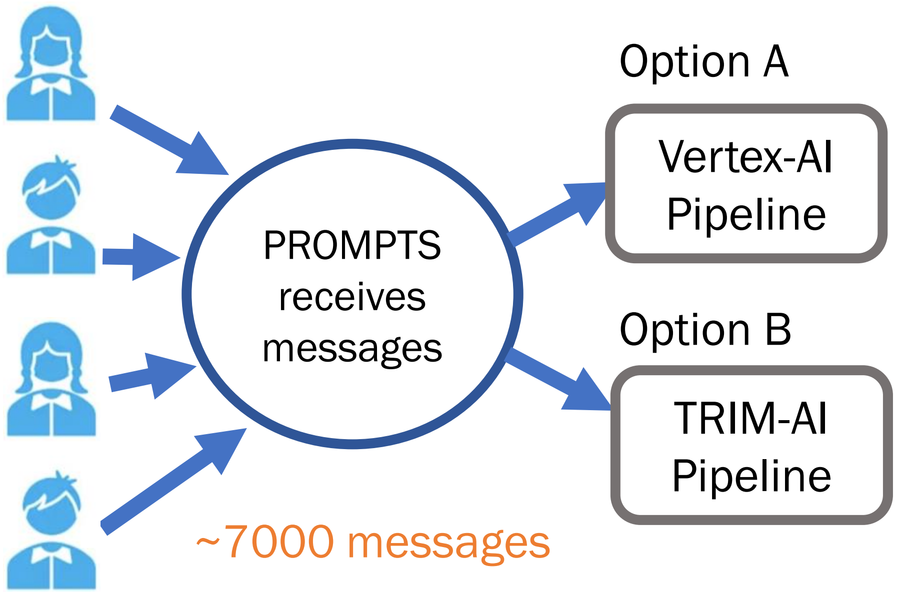
Performance Evaluation (A/B Test)



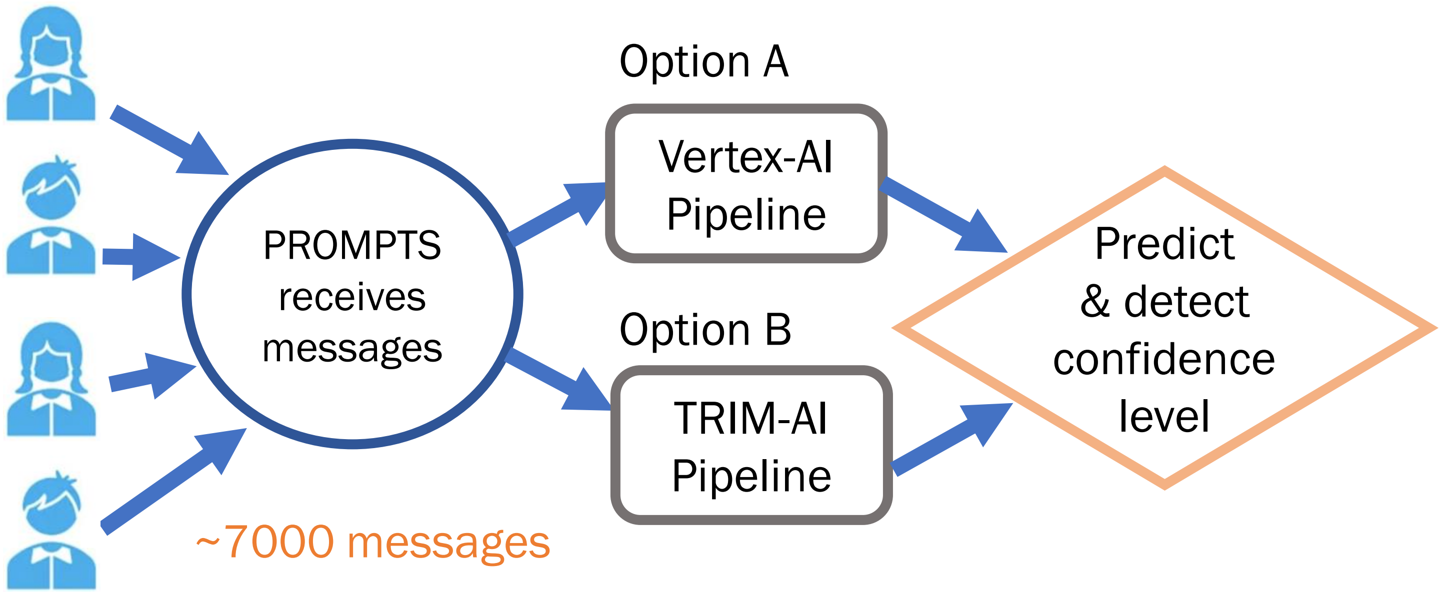
Performance Evaluation (A/B Test)



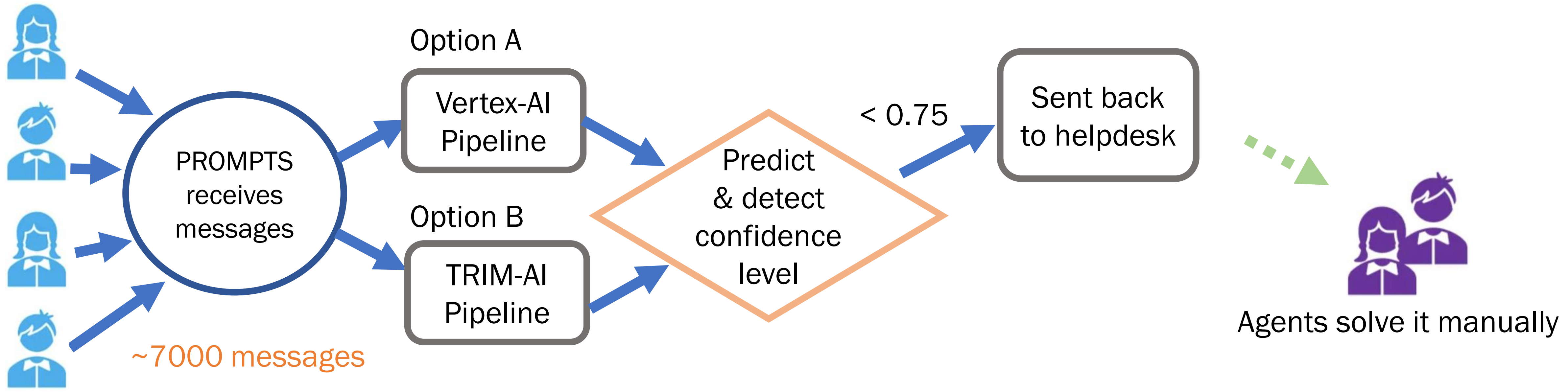
4 2 Performance Evaluation (A/B Test)



4 2 Performance Evaluation (A/B Test)

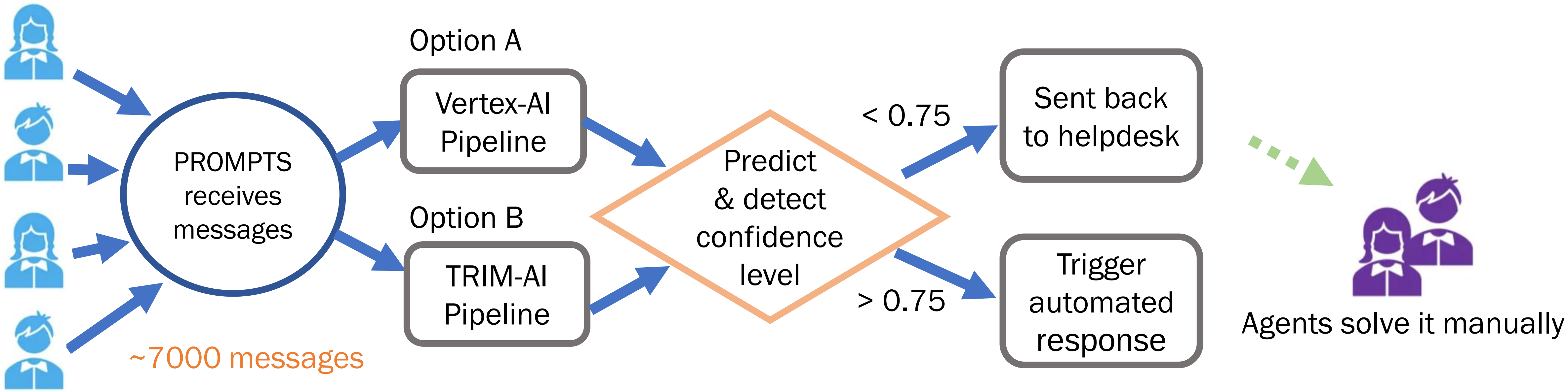


Performance Evaluation (A/B Test)

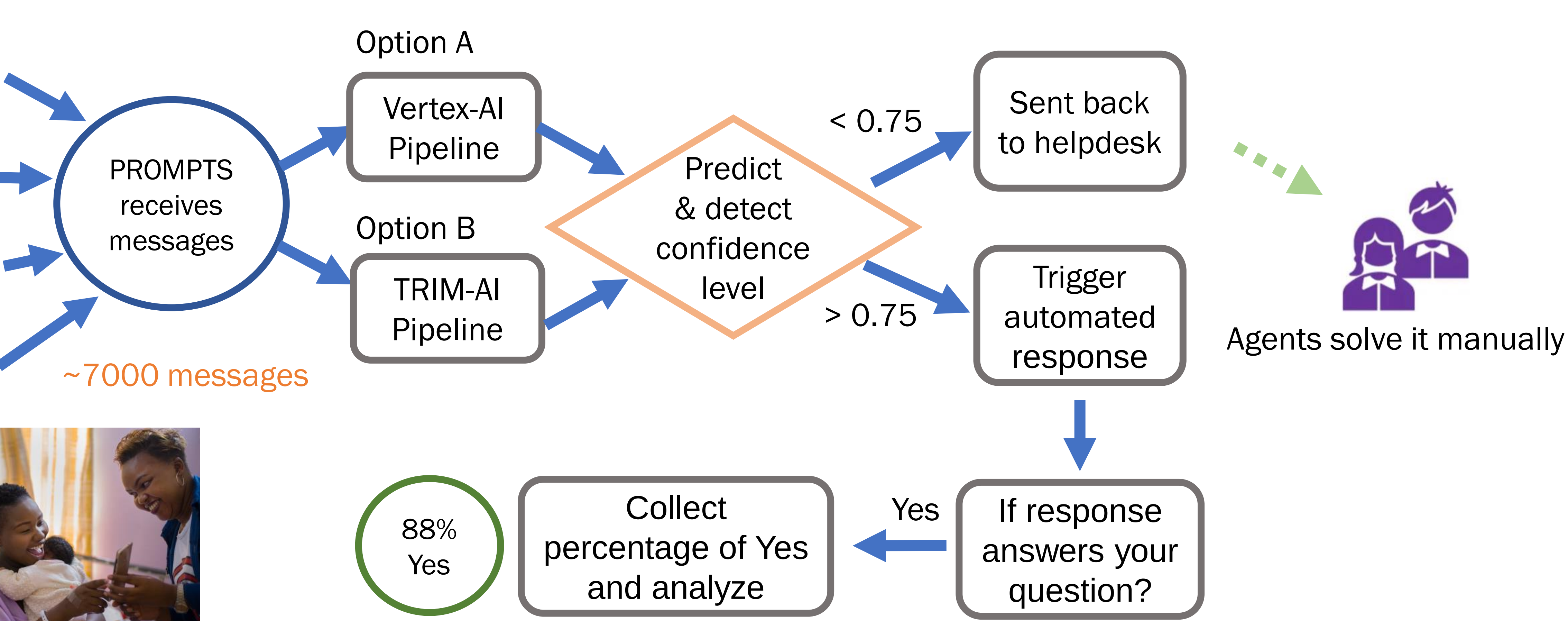


4 2

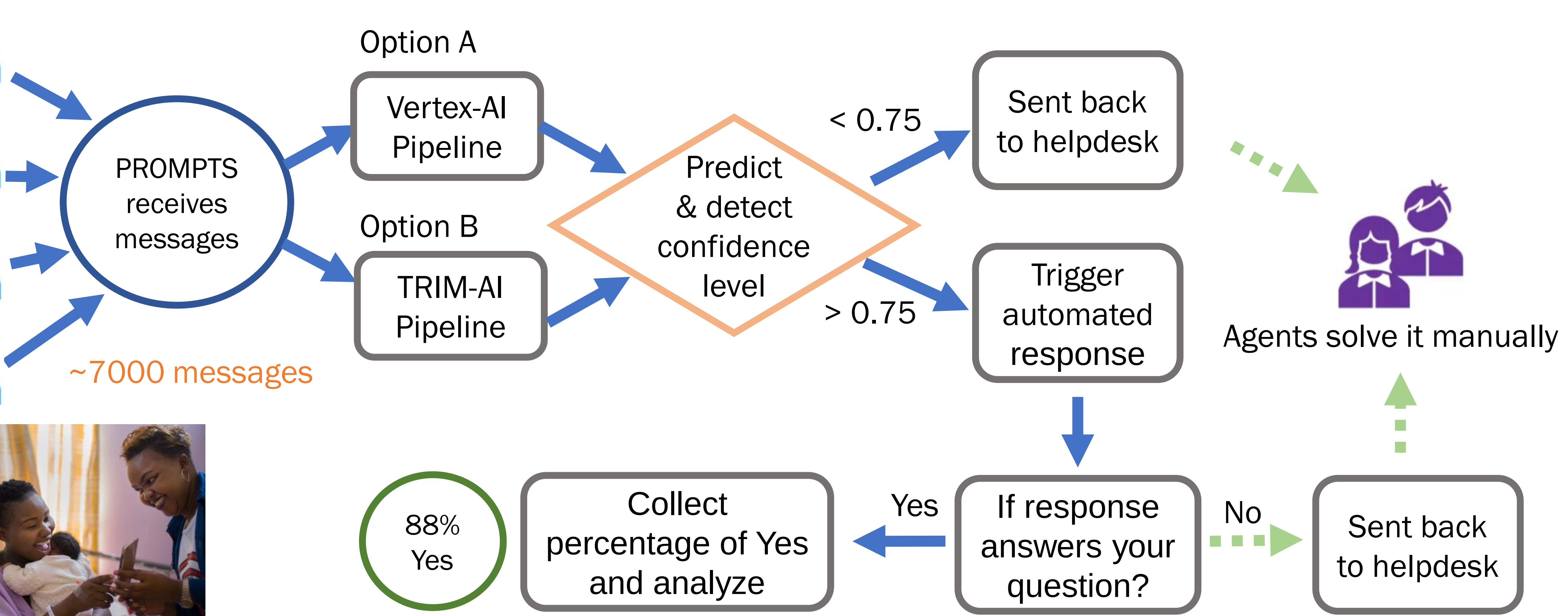
Performance Evaluation (A/B Test)



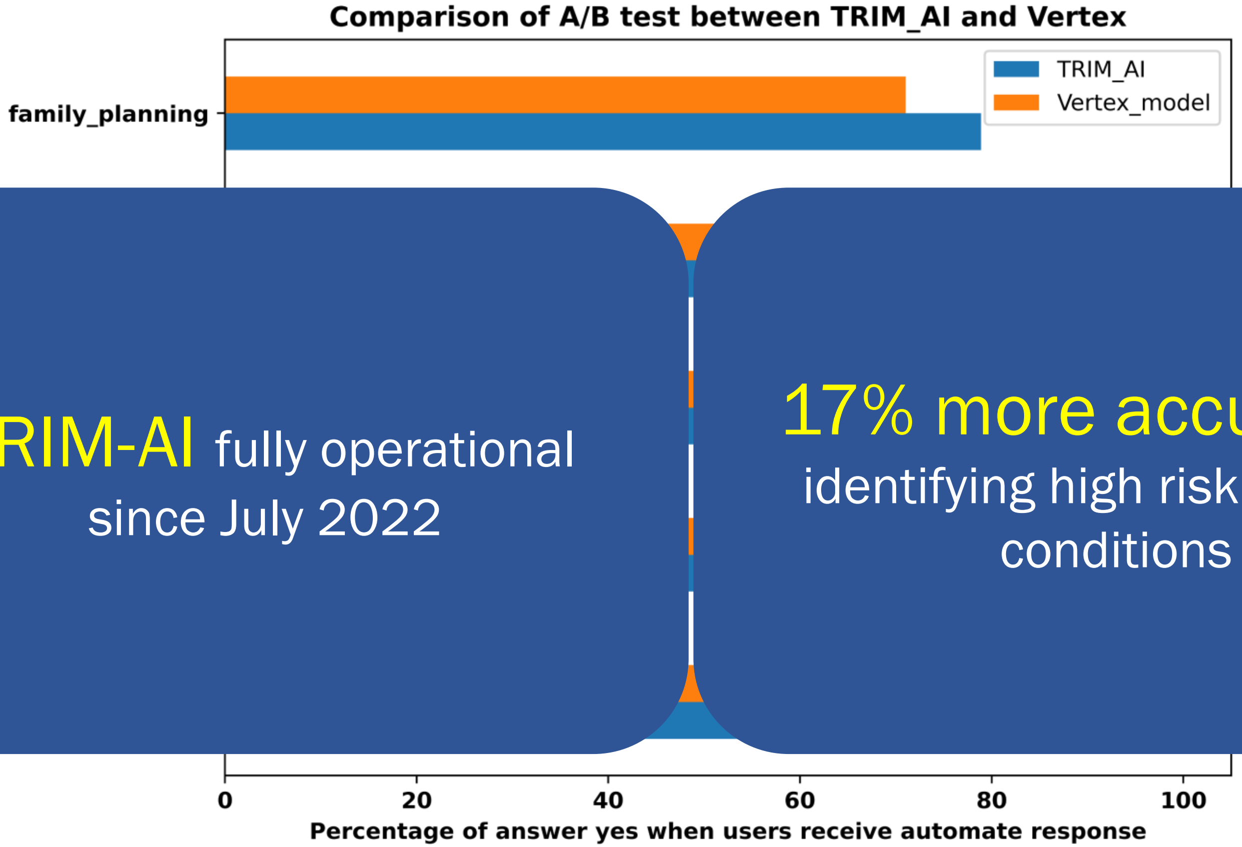
Performance Evaluation (A/B Test)



Performance Evaluation (A/B Test)



4 3 Performance Evaluation (A/B Test)

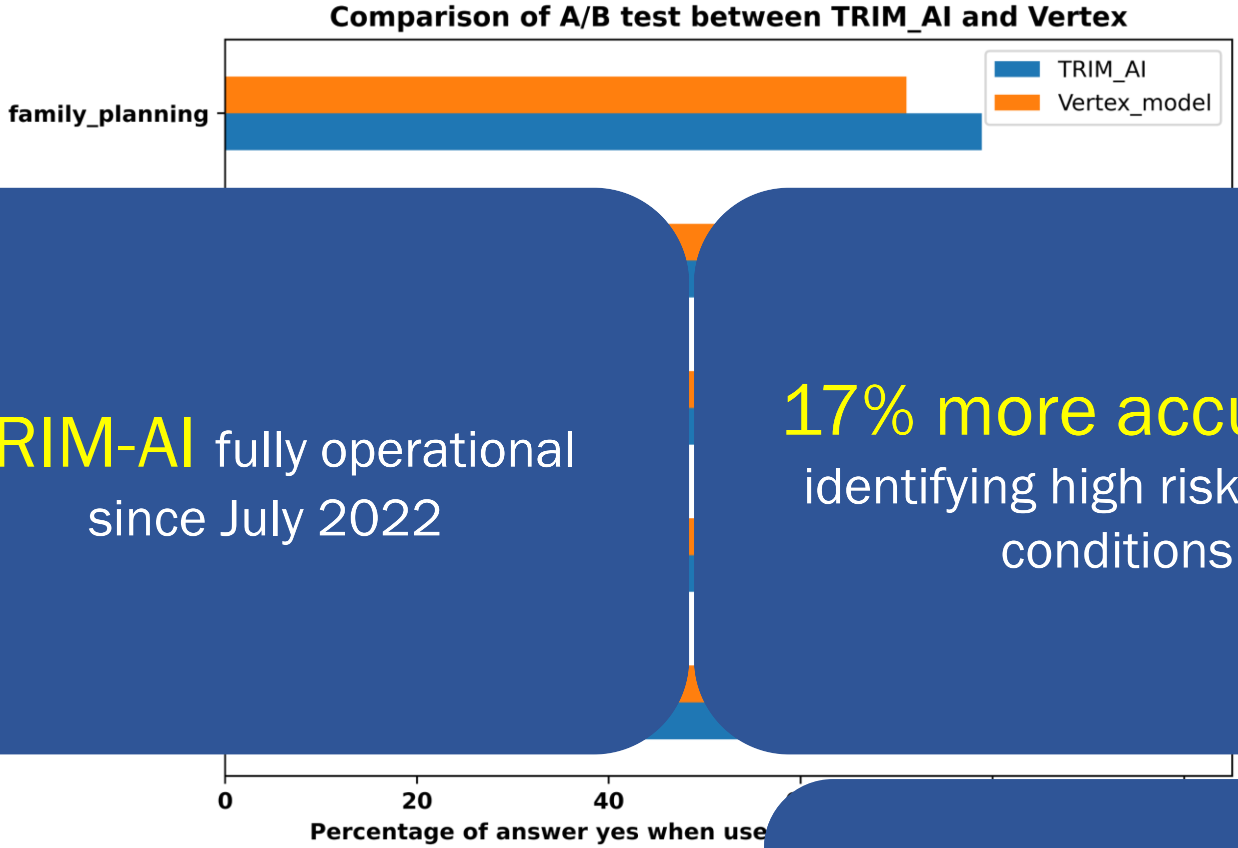


TRIM-AI fully operational since July 2022

17% more accurate at identifying high risk medical conditions

12% reduction in helpdesk workload

4 3 Performance Evaluation (A/B Test)



TRIM-AI fully operational since July 2022

17% more accurate at identifying high risk medical conditions

12% reduction in helpdesk workload

90% cost savings in model management costs

5 1

Summary

01 We propose TRIM-AI, an NLP-based framework for automated assessment of a pregnant woman's medical condition based on code-mixed messages

02 TRIM-AI utilizes multi-lingual pre-training and continual pre-training, achieves a weighted F1 score of 0.774

03 Real world A/B test shows that TRIM-AI is highly effective in generating high-quality predictions based on incoming SMS messages. Right now, TRIM-AI has been deployed inside PROMPTS and reduces the helpdesk workload by ~12%

Thank You & Questions?

Wenbo Zhang & Hangzhi Guo
(PhD Students)

Sathy Rajasekharan
(Exec Director, Jacaranda Health)

Jay Patel
(Tech Director, Jacaranda Health)

Beyond Buzzwords: Reimagining the Default Settings of Technology & Society

Ruha Benjamin, PhD
Professor, Department of African American Studies, Princeton University and Founding Director, Ida B. Wells Just Data Lab

Moderated by Karey M. Sutton, PhD
Scientific Director, Health Equity Research, MedStar Health Research Institute

AAMC CENTER FOR HEALTH JUSTICE

By Manzel Bowman

Beyond Buzzwords:

Reimagining the Default Settings of Technology & Society

Ruha Benjamin
Princeton University
Ida B. Wells Just Data Lab

AAMC Center for Health Justice
May 18, 2023

Art: Loveis Wise

Not all speed is movement.

-Toni Cade Bambara

two stories

new vision

racism distorts how we see
& how we are seen

Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites

Kelly M. Hoffman,[a,1] Sophie Trawalter,[a] Jordan R. Axt,[a] and M. Norman Oliver[b,c]

Global Phenomenon

The Mainichi

Japan's National Daily Since 1922

Search

Top Latest Japan World Business Sports Science Entertainment Opinion Lifestyle Obituaries

A thousand cuts: A 'Zainichi' Korean reporter's deep dive into microaggression in Japan

“Why I am in such demand
as a research subject,
when no one wants me
as a patient?”

–From People’s Science

The Washington Post
Democracy Dies in Darkness

Health

Racial bias in a medical algorithm favors white patients over sicker black patients

MOTHERBOARD
TECH BY VICE

'Significant Racial Bias' Found in National Healthcare Algorithm Affecting Millions of People

the new jim code

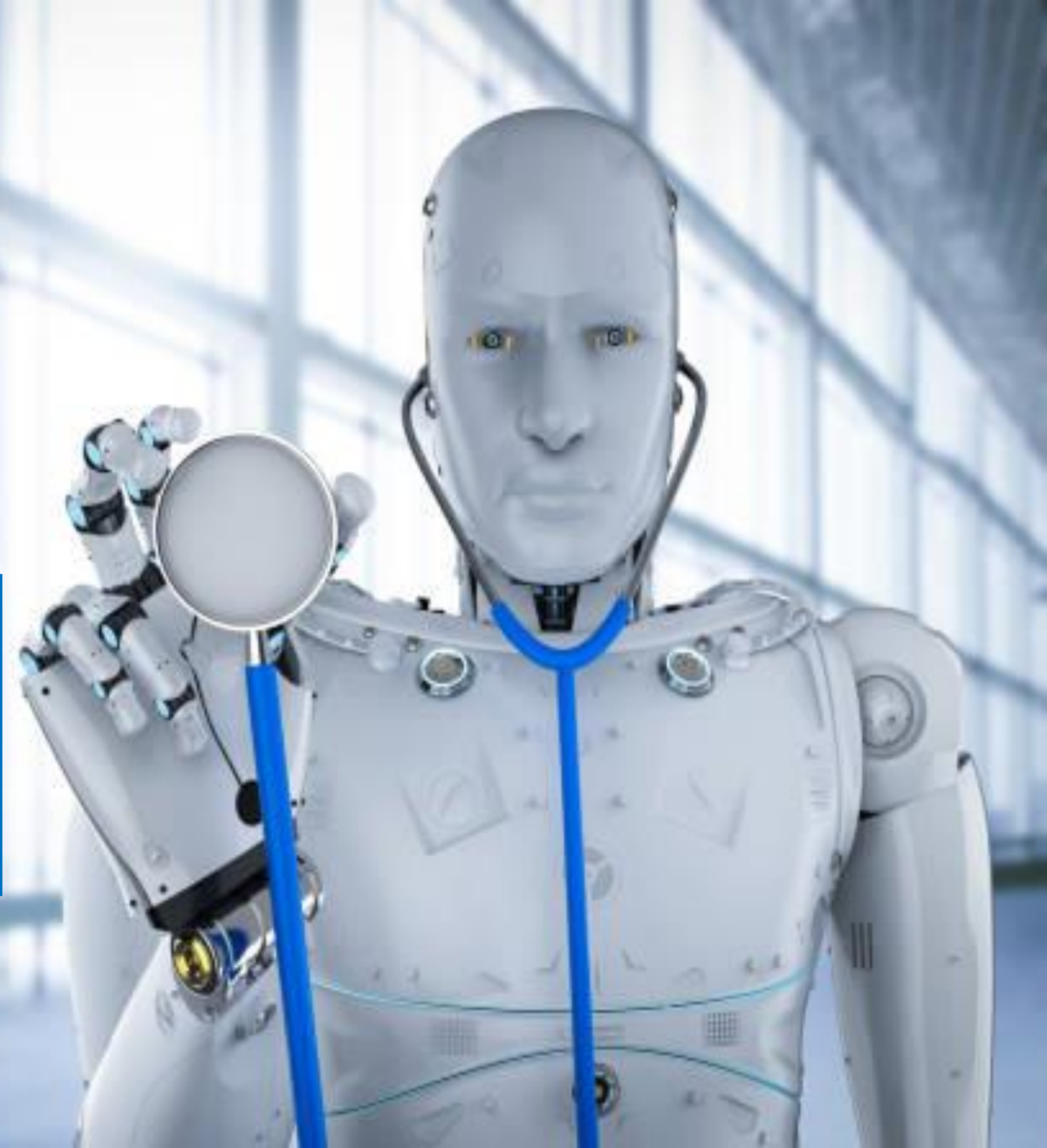
; coded bias + imagined objectivity

; innovation that enables containment

deeeeep learning

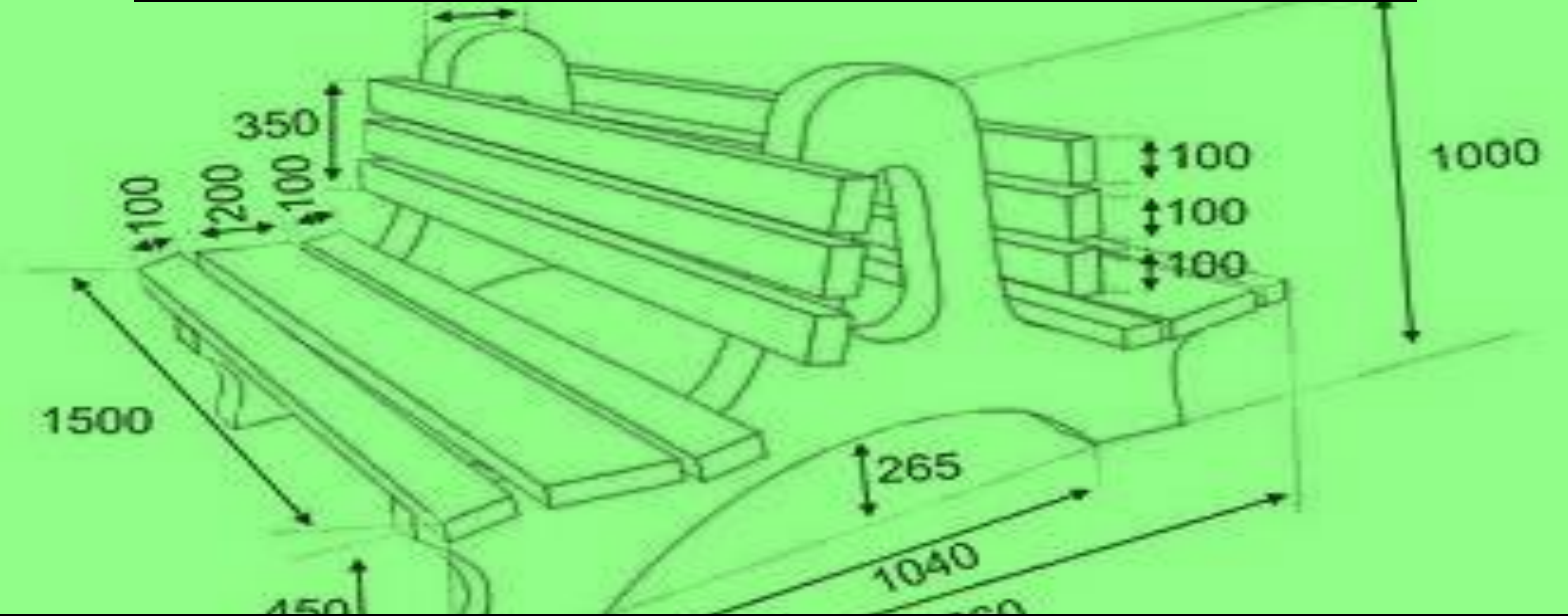
Computational depth without historical or sociological depth?

SUPERFICIAL LEARNING



encoded inequity

discriminatory design



Spikes in digital tools?

obvious vs insidious

obvious vs insidious

Algorithmic Accountability Act

NEWS / PRESS RELEASES

February 03, 2022

Wyden, Booker and Clarke Introduce Algorithmic Accountability Act of 2022 To Require New Transparency And Accountability For Automated Decision Systems

Legislation Requires Assessment of Critical Algorithms and New Public Disclosures; Bill Endorsed by AI Experts and Advocates; Bill Will Set the Stage For Future Oversight by Agencies and Lawmakers

BLUEPRINT FOR AN

AI BILL OF RIGHTS

MAKING AUTOMATED SYSTEMS WORK FOR THE AMERICAN PEOPLE

OCTOBER 2022

THE WHITE HOUSE
WASHINGTON

HEALTH TECH

STAT+

Could training AI to learn from patients give doctors a better understanding of the causes of pain?

By Erin Brodwin Feb. 2, 2021

Reprints

MIKE REDDY FOR STAT

- 36,369 observations from 4,172 patients.

“We didn't train the algorithm to predict what the doctor was going to say about the X-ray.

We trained it to predict what the patient was going to say about their own experience of pain in the knee."

final proposition

if inequity is woven
into the very fabric of society
then each twist, coil, and code
is a chance for us to weave
new patterns, practices, politics
its vastness will be its undoing
once we accept that we are

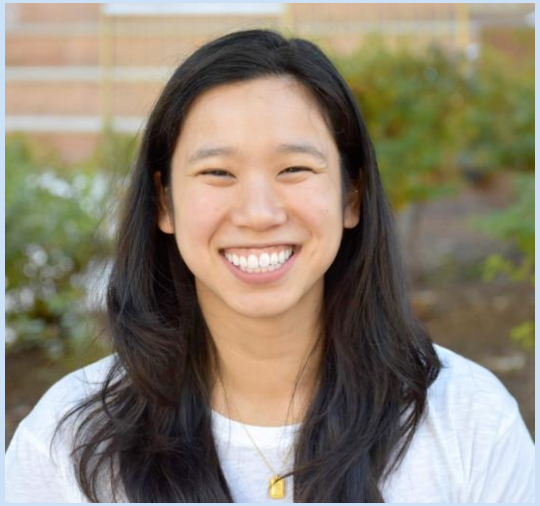
pattern makers.

Crowdsourcing Exercise: Designing Principles for NLP in Maternal Health

AAMC Center for Health Justice
&
Allen Institute for AI

Maria Antoniak, PhD
Allen Institute for AI
NLP, bias, reproductive healthcare

Carla S. Alvarado, PhD, MPH
AAMC Center for Health Justice
health policy, public health, intersectionality, health equity



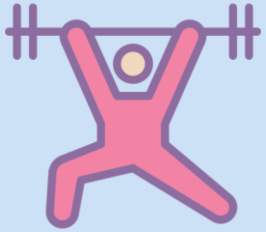
Lucy Lu Wang, PhD
University of Washington
health informatics, NLP, accessibility

Irene Chen, PhD
UC Berkeley, UCSF
ethics of machine learning for health

Aakanksha Naik, PhD
Allen Institute for AI
medical NLP, understudied language

Our Team

Assisted by an amazing team of volunteer facilitators!



Affrille Degoma

Kathryn Brand

Jennifer Bretsch

Keith Krosinky

Dallas Peoples

Tracey Alcendor Robinson

Mary Heller

Mobie Nwaokomah

Mubarak Childs

Sherrie Reece

Ebonie Megibow

Our goal: guiding principles for NLP & maternal health

Your expertise and lived experience is invaluable.

We want to hear your hopes, fears, goals, critiques — all of it!

We share many of your concerns about the use of AI and NLP, and we want to build the best possible future, together with you.

What to expect from this session

- An interactive exercise with a chatbot
- A moderated discussion about your experiences and opinions about NLP and maternal health
- A short survey asking for your perspectives

From Data to Action: What Public Health, Hospitals and Health Systems Can Do

Simon Linwood, MD, MBA
Janette Robinson Flint
Justin Schonfeld, PhD
Carlos Siordia, PhD, MS

Moderated by Karey M. Sutton, PhD
Scientific Director, Health Equity Research, MedStar Health Research Institute

Meet the Panelists

Simon Linwood, MD, MBA
Chief Information Officer and Chief Information Officer, University of California Riverside Health and School of Medicine

Janette Robinson Flint
Executive Director, Black Women for Wellness

Justin Schonfeld, PhD
Research Scientist, Public Health Agency of Canada

Carlos Siordia, PhD, MS
Lead Interdisciplinary Epidemiologist, Division of Violence Prevention, National Center for Injury Prevention and Control

Drawing Change Workshop Summary

Yolanda Liman
Drawing Change

AAMC CENTER FOR HEALTH JUSTICE

MATERNAL HEALTH EQUITY WORKSHOP:
FROM STORY TO DATA TO ACTION
MAY 18, 2023

Thank you for joining us!

Association of
American Medical Colleges

Let’s Keep in Touch

Please use the hashtags:
#healthequity and #maternalhealth
And don't forget to tag us: @AAMCjustice

aamchealthjustice.org

healthjustice@aamc.org

@AAMCjustice

Sign up for the Center for Health Justice Newsletter