COVID-19 Public Policy Conformity Analysis

Mohammad Sabik Irbaz
Former Lead ML Engineer,
Omdena Inc. and Pioneer Alpha Ltd.
sabikirbaz@iut-dhaka.edu

Dr. Md. Ruhul Amin
Assistant Professor, Department of CIS,
Fordham University
mamin17@fordham.edu

Rupak Sarkar
PhD Student, Department of CS,
University of Maryland, College Park
rupak@cs.umd.edu

Dr. Ashiqur R KhudaBukhsh
Assistant Professor, Department of CIS,
Rochester Institute of Technology
akhudabu@cs.cmu.edu

**Fordham Human Centered AI Research (FAIR) Lab,
Fordham University**

Research Objective

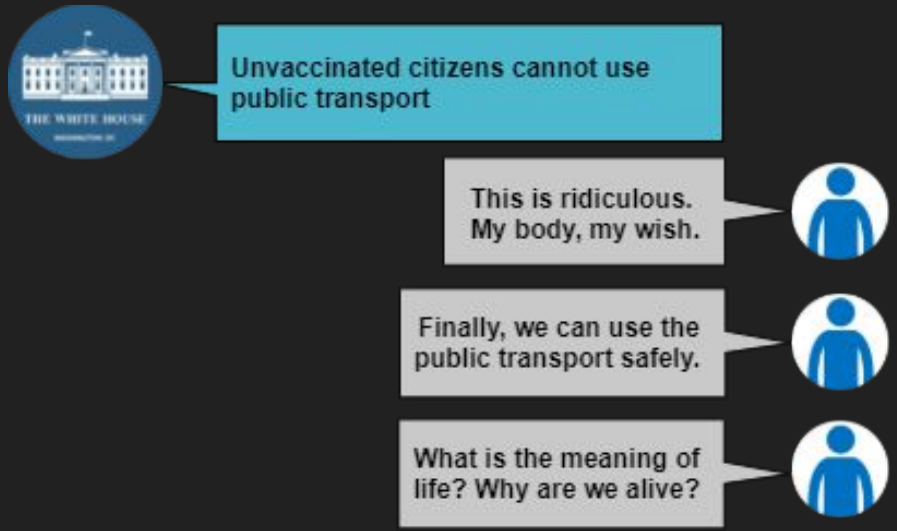
The main objective of this research is to:

1) analyze the public stance on government policies across the world during the Covid-19 pandemic

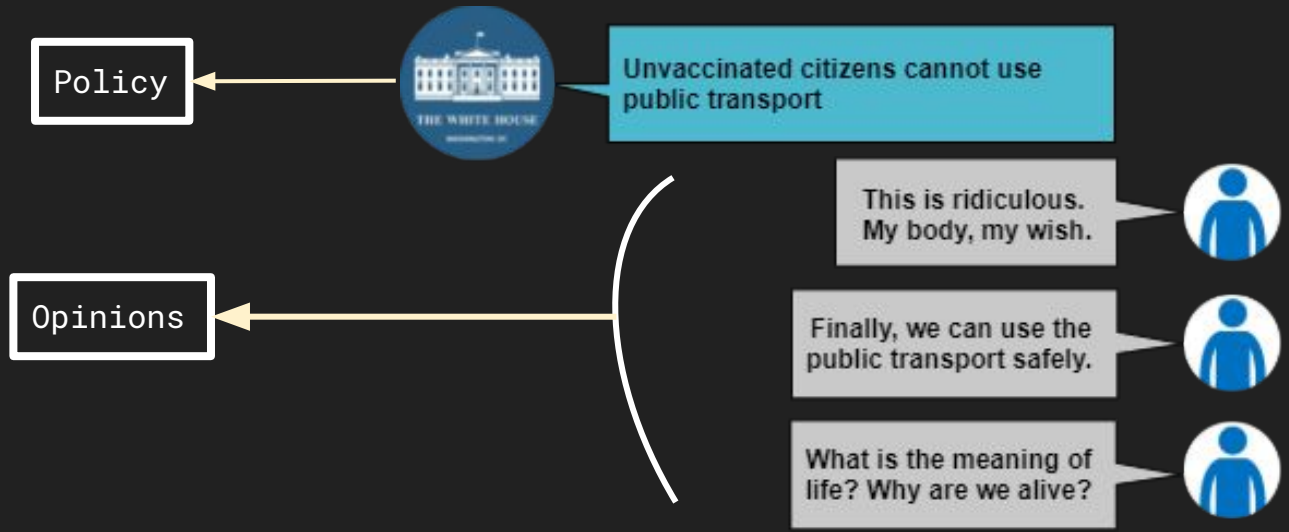
2) correlate with real-life implications of those aggregate stances using historical data.

3) build a reliable tool for policy-makers to help take effective measures during similar situations in the future.

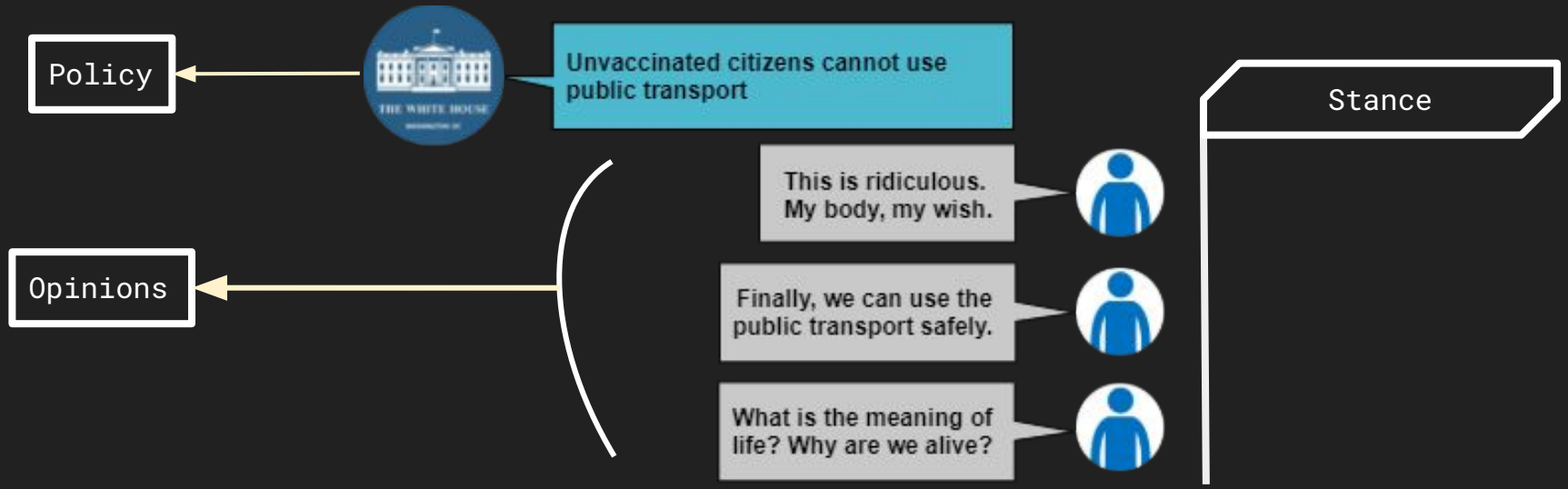
Objective 01



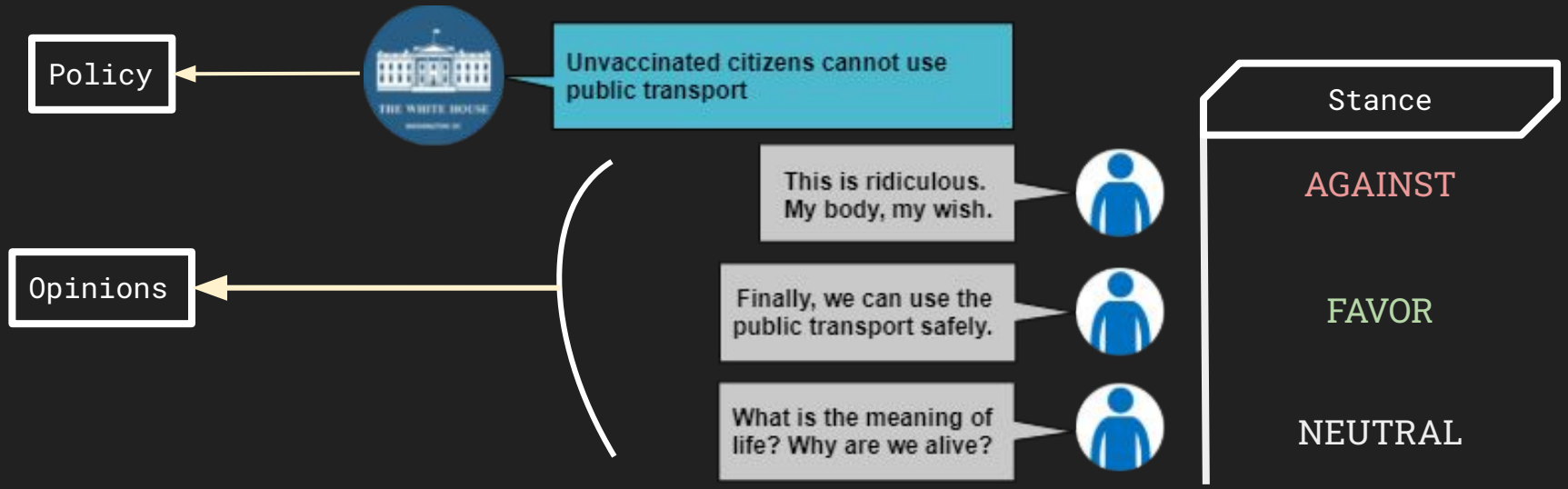
Objective 01



Objective 01



Objective 01



Dataset

OxCGRT (Oxford Covid-19 Government Response Tracker)

OxCGRT collects publicly available information on 21 indicators of government policies for COVID-19.

C1 - C8 => Containment and Closure Policies
E1 - E4 => Economic Policies
H1 - H8 => Health System Policies
M1 => Miscellaneous Policies

The dataset is organized in a way that we can get country and state level data.

It is still being updated every hour as we speak.

They also provide a time-series data from where we can interpret the longevity and stringency of any policy.

Dataset (cont.)

COVID-19 related Tweets

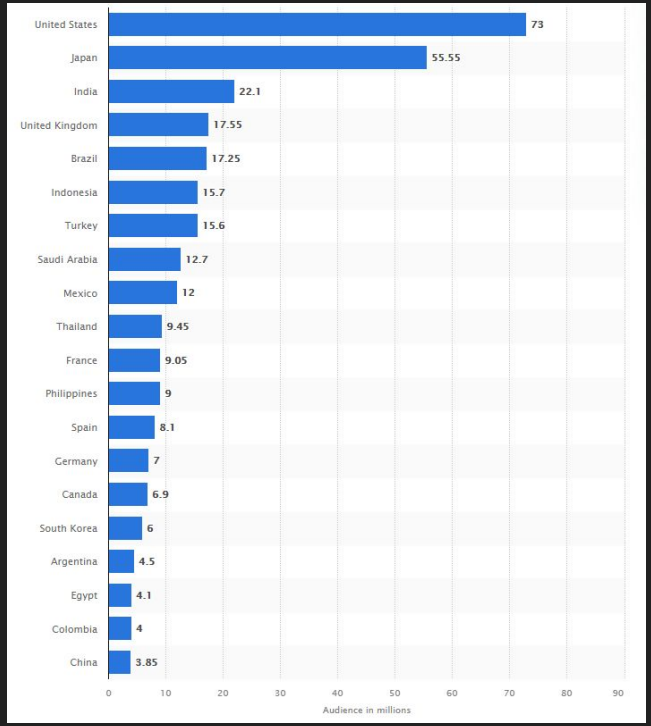
In one of the previous works, FAIR Lab gathered more than 1B COVID-19 related tweets from April 2020 onwards.

Dataset (cont.)

COVID-19 related Tweets

In one of the previous works, FAIR Lab gathered more than 1B COVID-19 related tweets from April 2020 onwards.

A diverse group of people from different countries use Twitter.



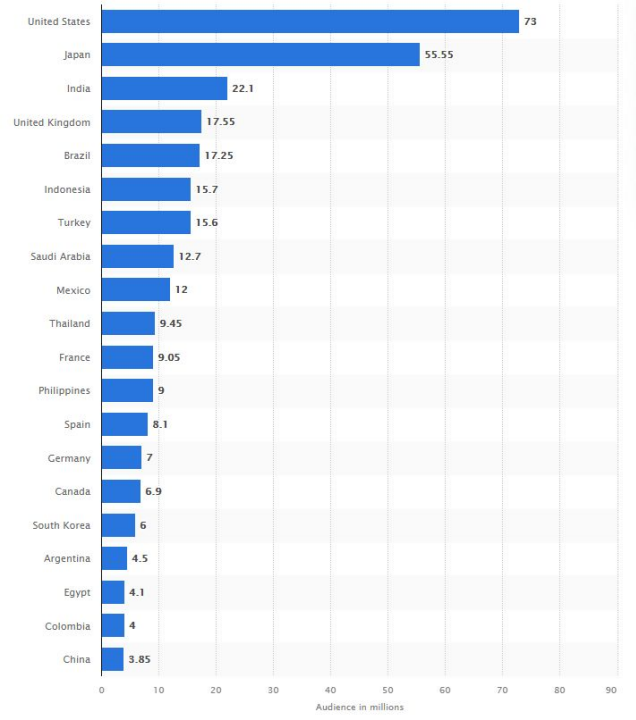
Source: Twitter 2021 Official Letter to Shareholders

Dataset (cont.)

COVID-19 related Tweets

In one of the previous works, FAIR Lab gathered more than 1B COVID-19 related tweets from April 2020 onwards.

A diverse group of people from different countries use Twitter.



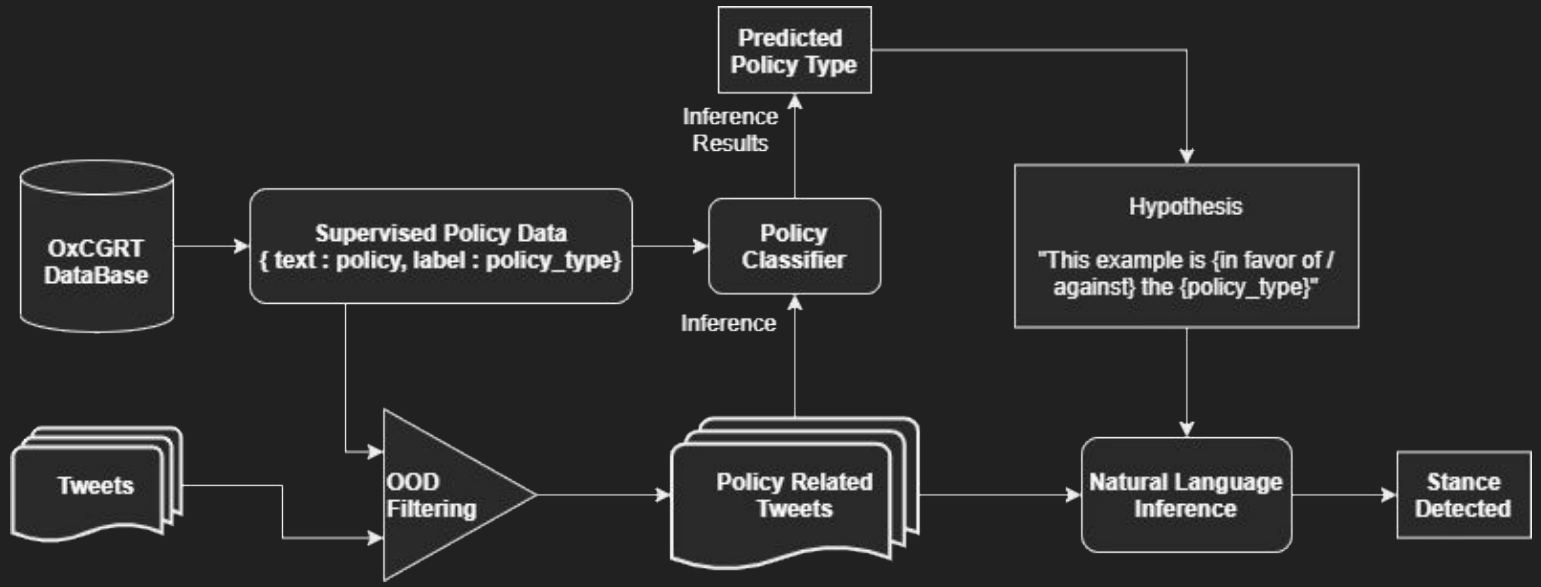
Why are we using Twitter Data ?

1) Easily extractable with hashtags and proper queries.

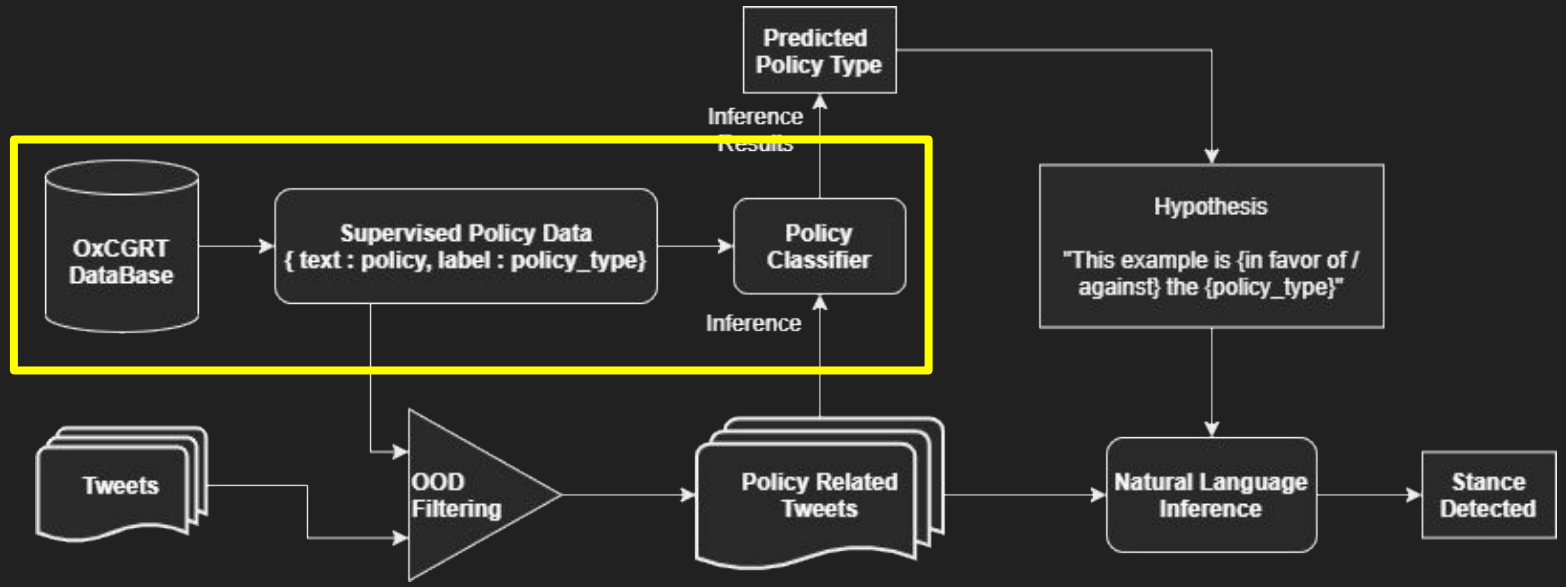
2) Twitter Stats:
- Twitter has 199 million daily active users.
- 500 million tweets per day
- Top social network platform in Japan
- 6th ranked mobile app
- 42% of US Twitter users are between the ages of 18 and 29, 27% of users are 30-49 years old.

Source: Twitter 2021 Official Letter to Shareholders

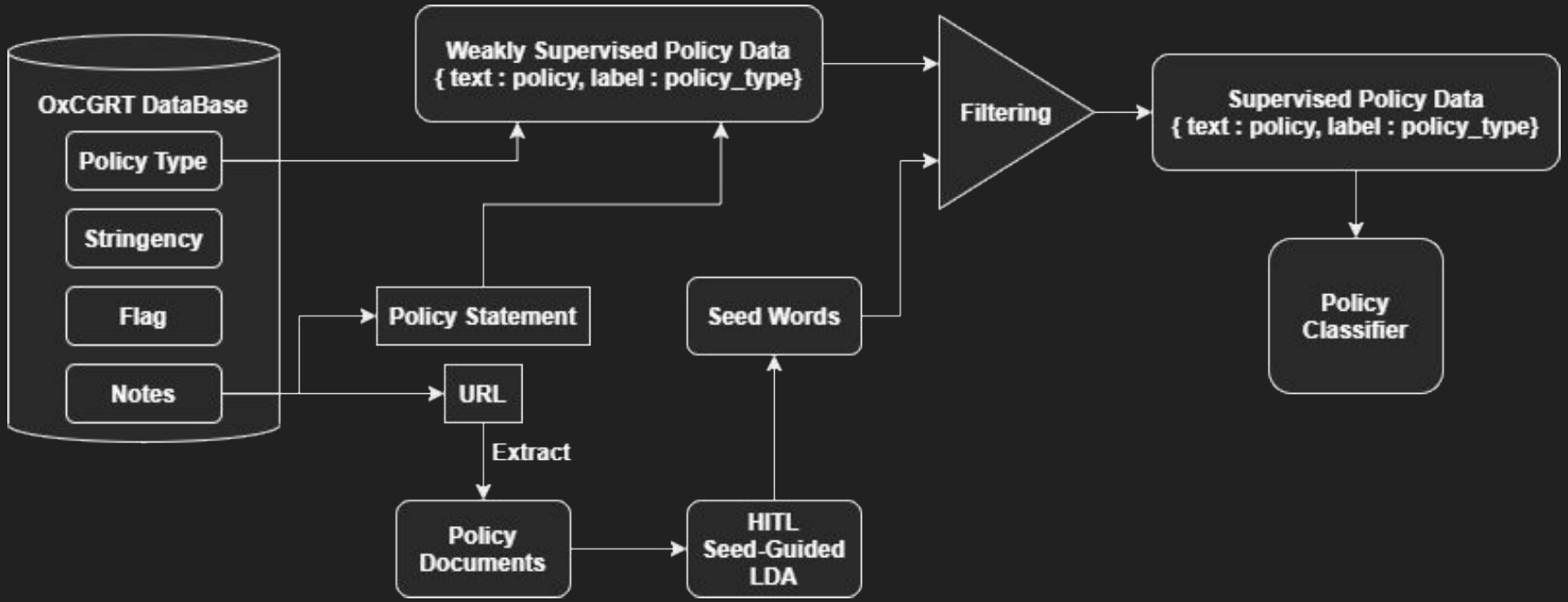
Workflow (High-level Overview)



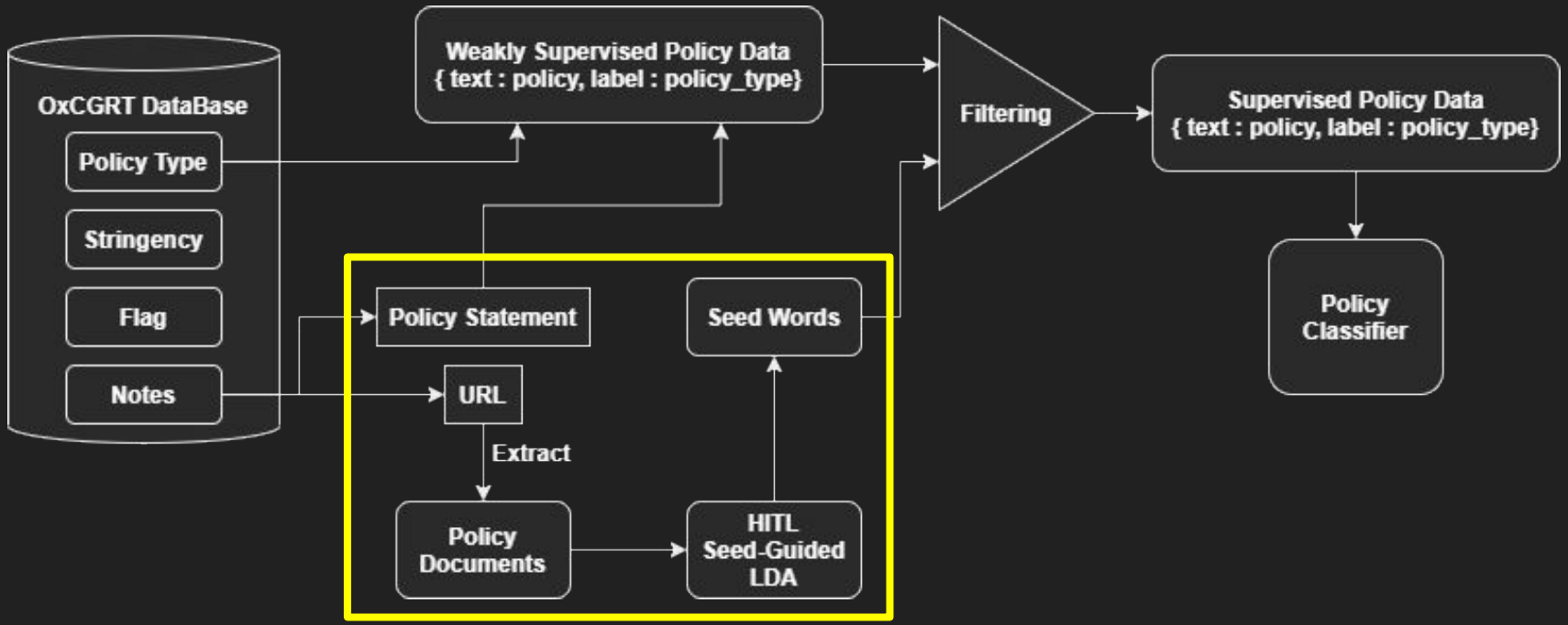
Workflow (High-level Overview)



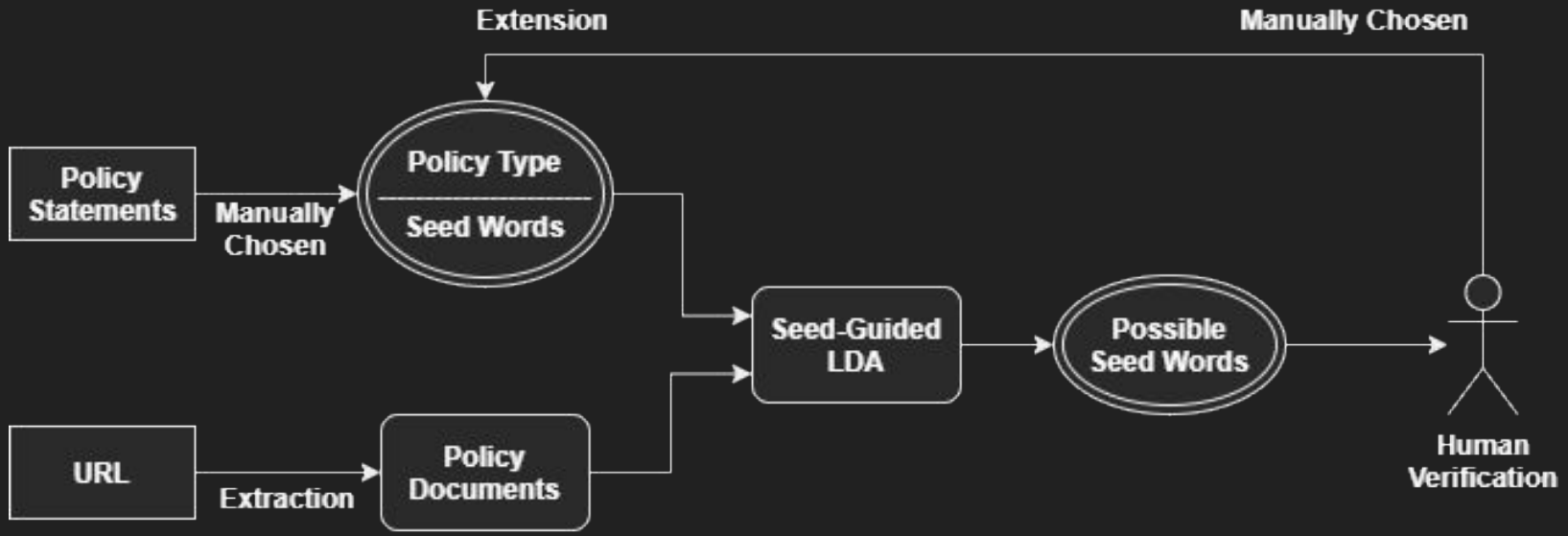
Workflow (Policy Classifier)



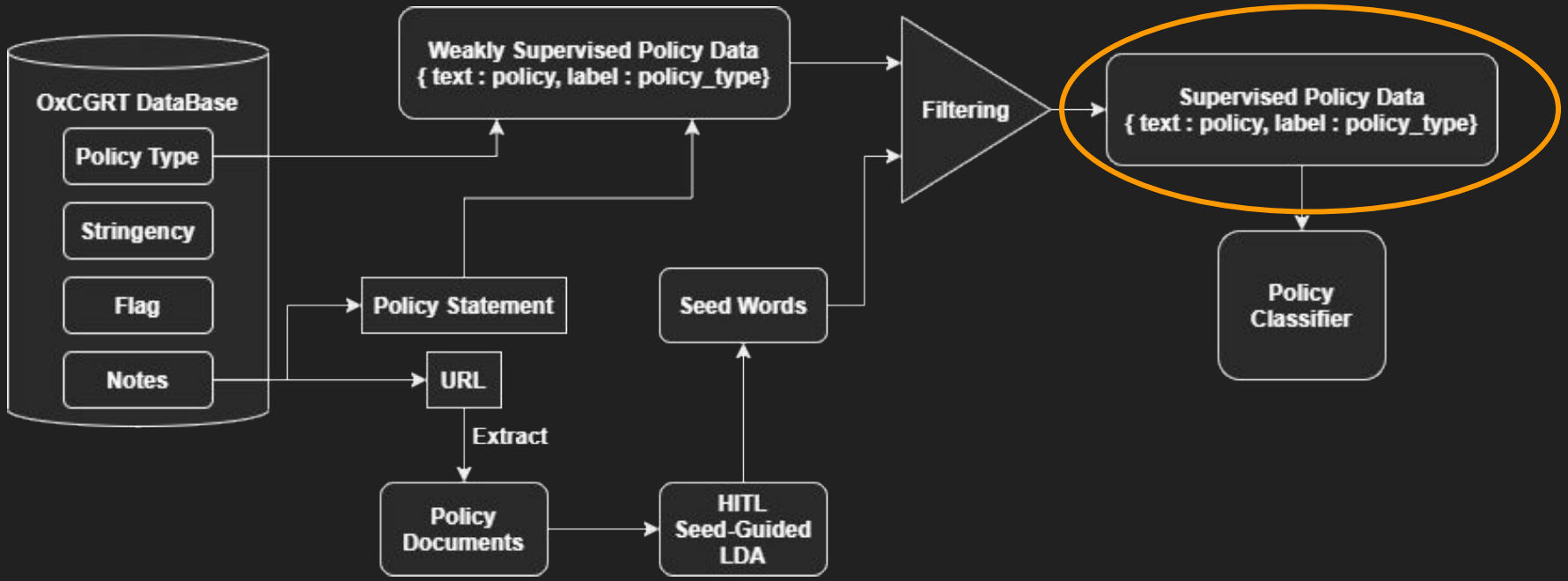
Workflow (Policy Classifier)



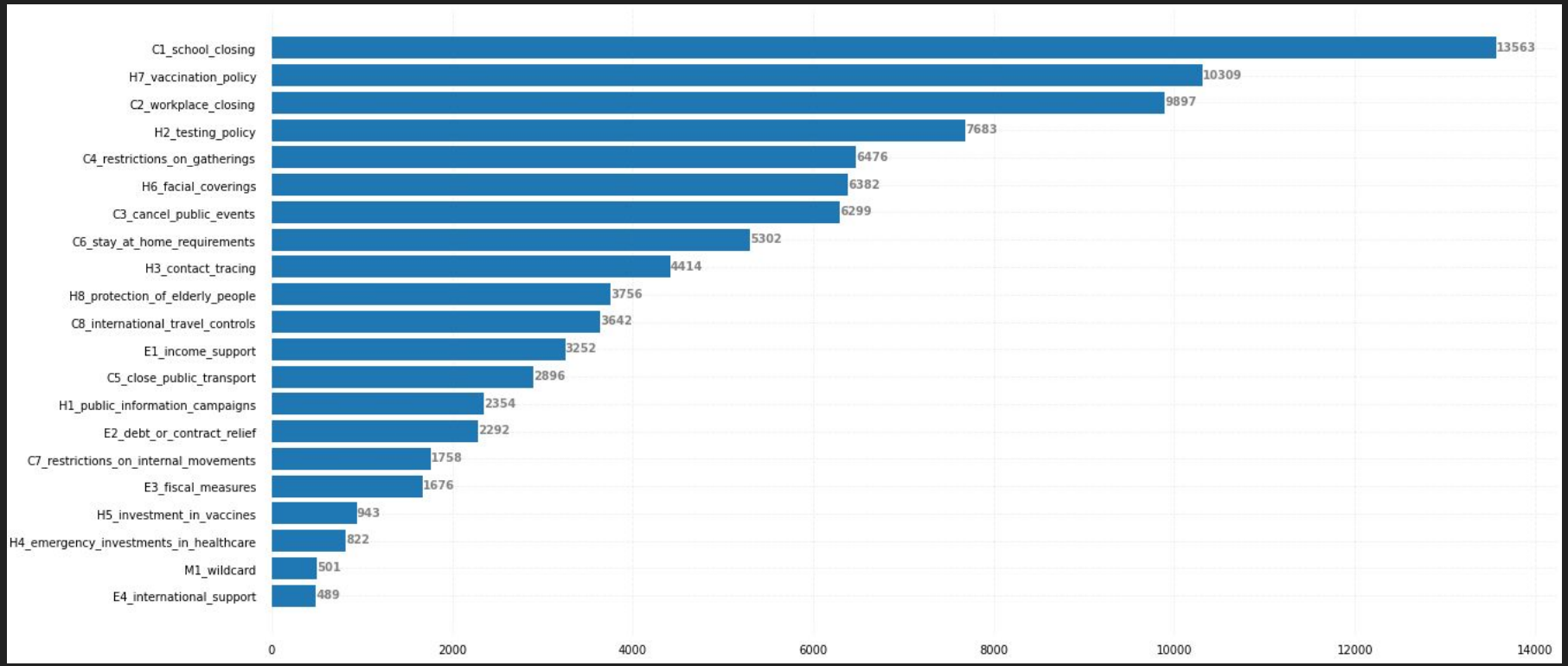
Workflow (HITL Seed-Guided LDA)



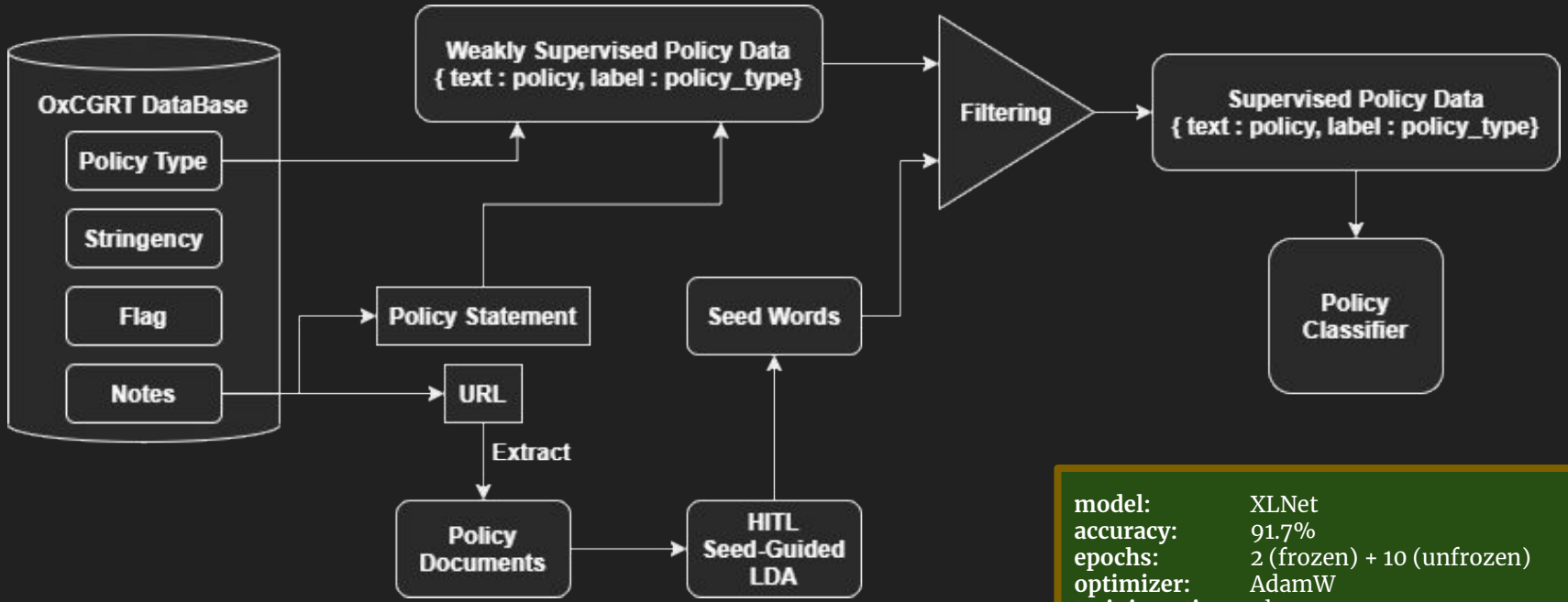
Workflow (Policy Classifier)



Policy Data Distribution

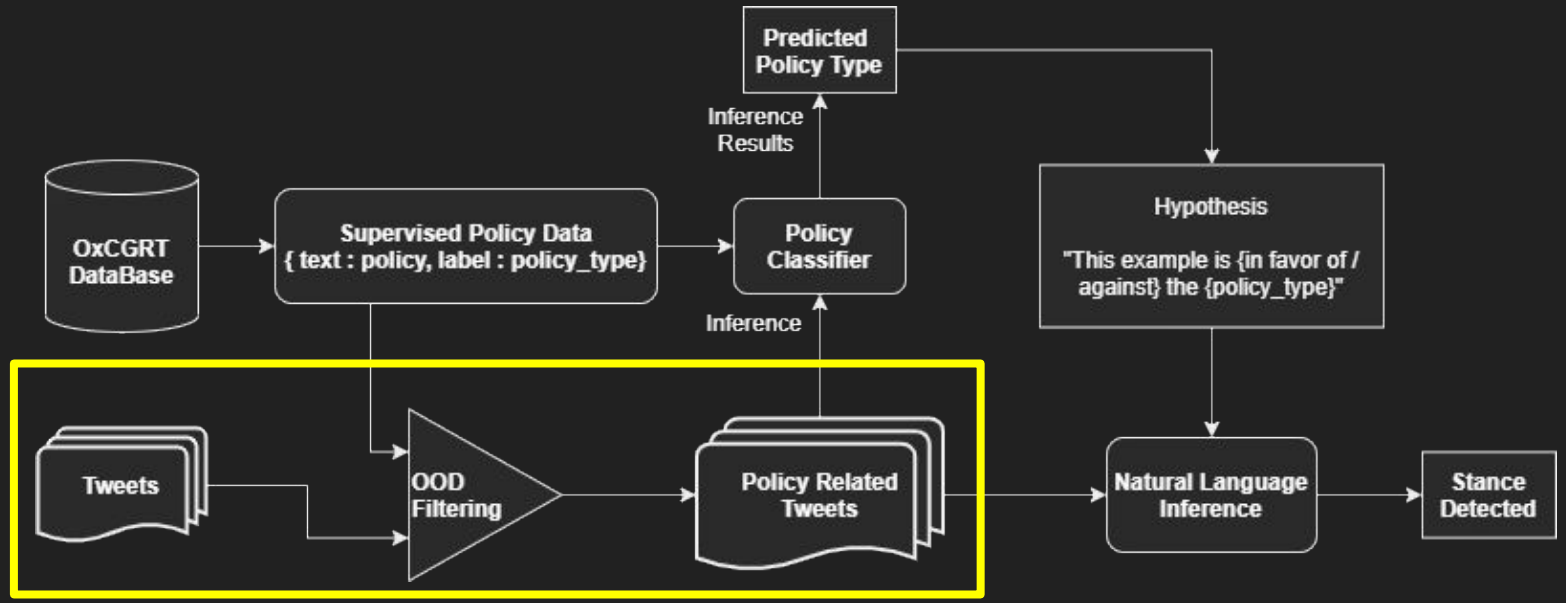


Workflow (Policy Classifier)

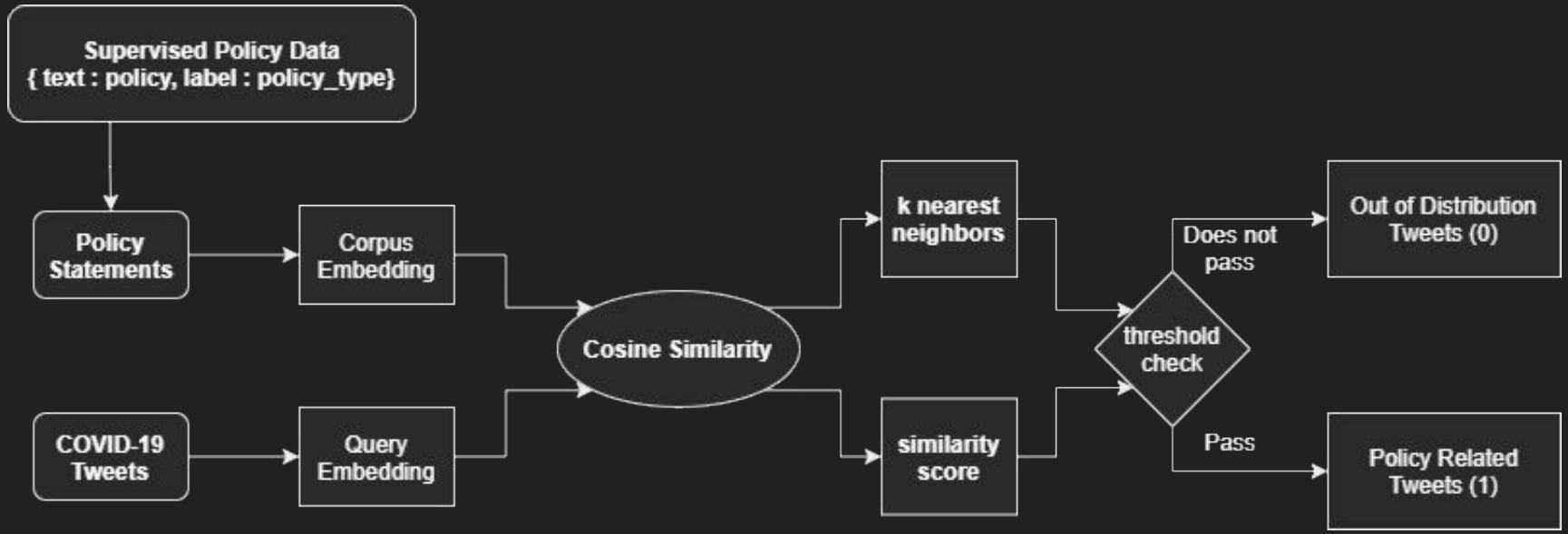


model:	XLNet
accuracy:	91.7%
epochs:	2 (frozen) + 10 (unfrozen)
optimizer:	AdamW
training_time:	5 hours

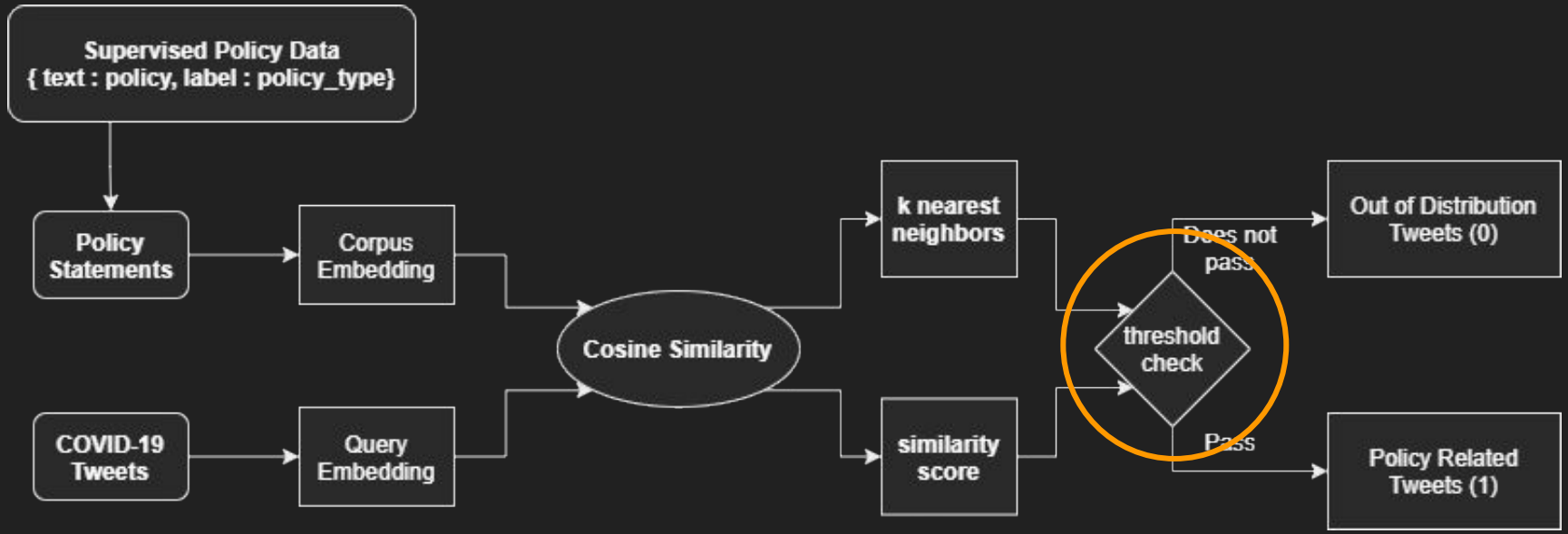
Workflow (High-level Overview)



Workflow (Filtering Unrelated Tweets)



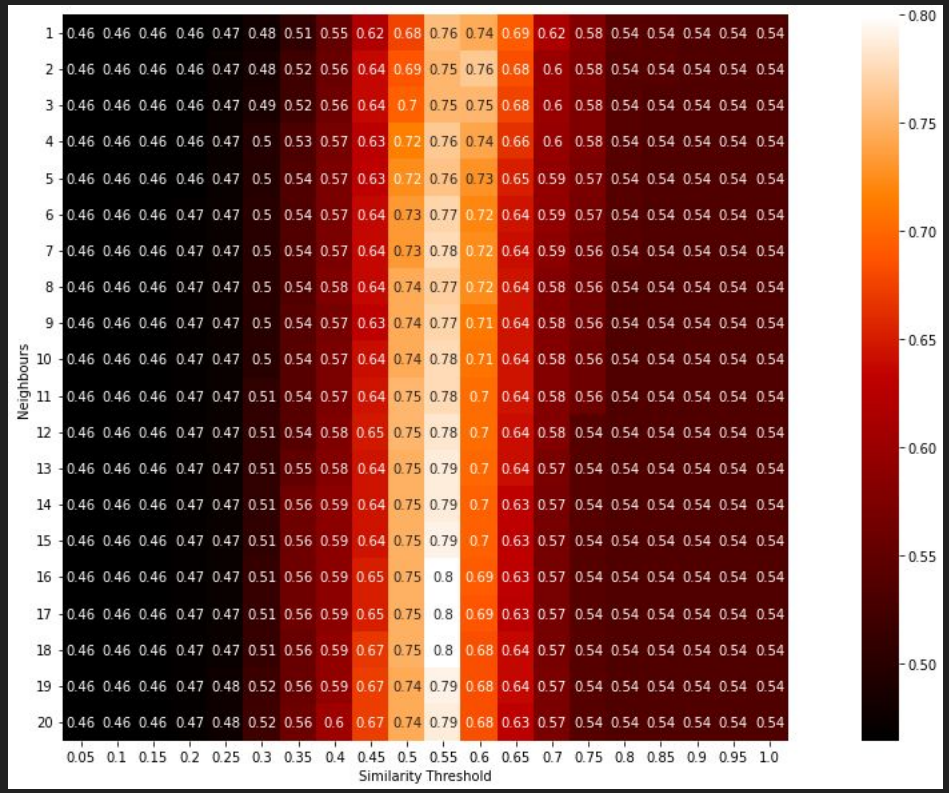
Workflow (Filtering Unrelated Tweets)



Hyperparameter Tuning

Manually Annotated Data For OOD Filtering

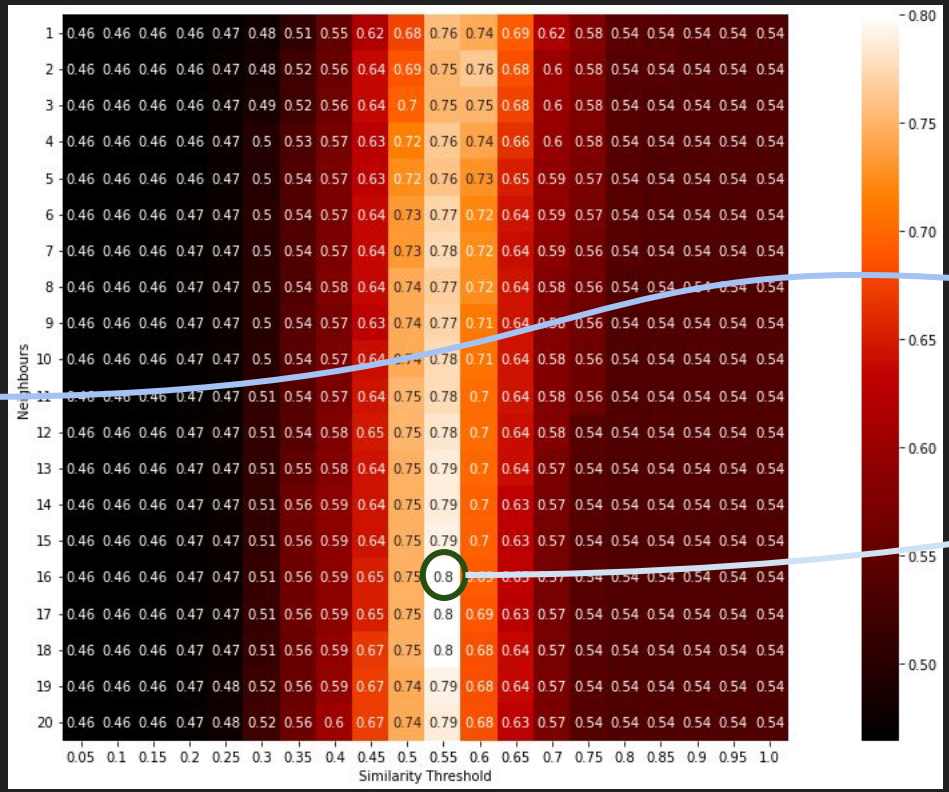
{text: tweet,
ood_label: 0 or 1}



Hyperparameter Tuning

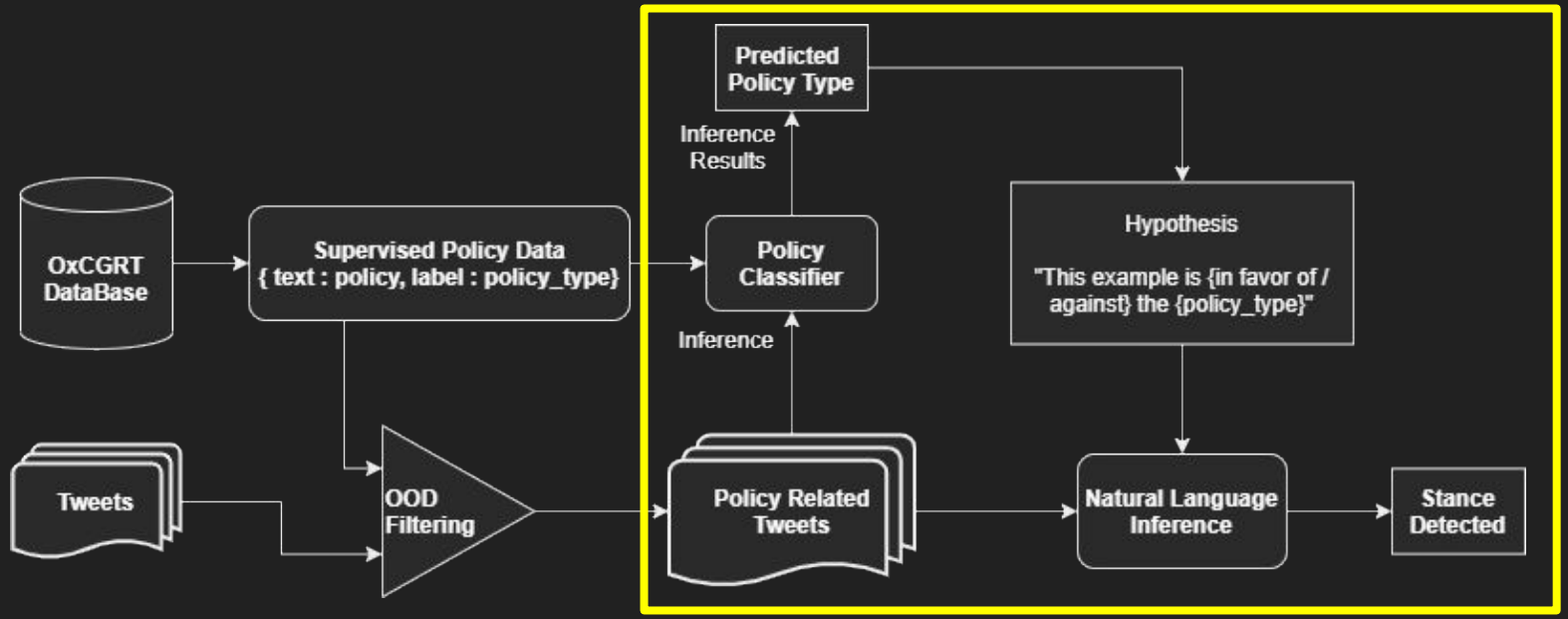
Manually Annotated Data For OOD Filtering

{text: tweet,
ood_label: 0 or 1}

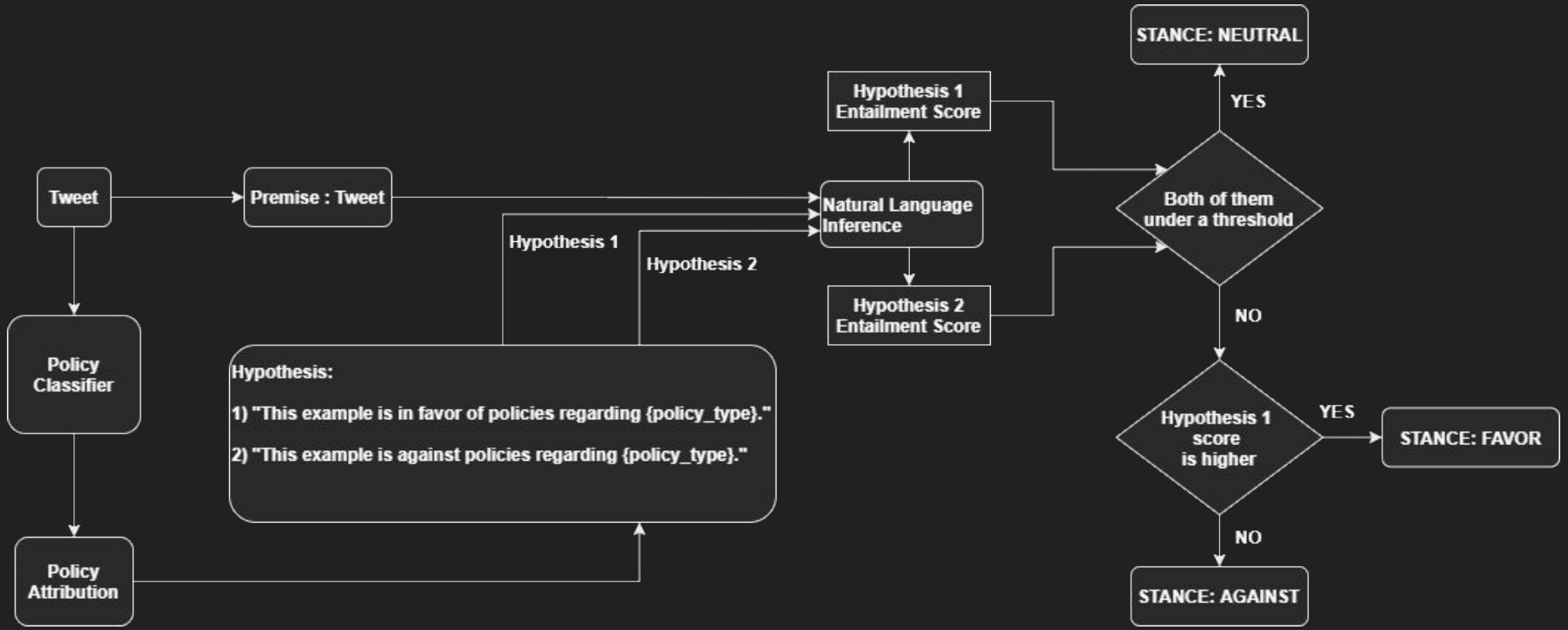


K : 16
Sim_thresh : 0.55
Accuracy : 80%

Workflow (High-level Overview)



Workflow (Stance Detection)



Contribution (till now)

Supervised Data Collection, Cleaning and Processing for Public Policy Type Classification.

Out of Distribution Data Filtering and Manual Annotation for OOD Filter Test.

Using the NLI concept in Stance Detection Problem.

Future Works

Building Test Data to validate the Stance Detection Workflow

Correlate the implications of the aggregated stance scores using historical data.

Building a reliable tool for policy-makers to help take effective measures during similar situations in the future.