

Master Thesis Defense Presentation

DESIGN AND IMPLEMENTATION OF SMART CHATBOT USING DEEP LEARNING

Rianna Dsilva [11010632]

SRH Hochschule Heidelberg - Department of Information Technology
Under the Guidance of Sofia Trojanowska and Dr. Achim Gottscheber

Today's Agenda

Introduction

Problem Statement & Objective

Architecture of Designed Model

Methodology

Software Implementation

Results

Challenges in Development

Conclusion

TensorFlow and Thesis Timeline

Questions

Introduction

"A Conversational Agent is a computer program written to participate in conversation."

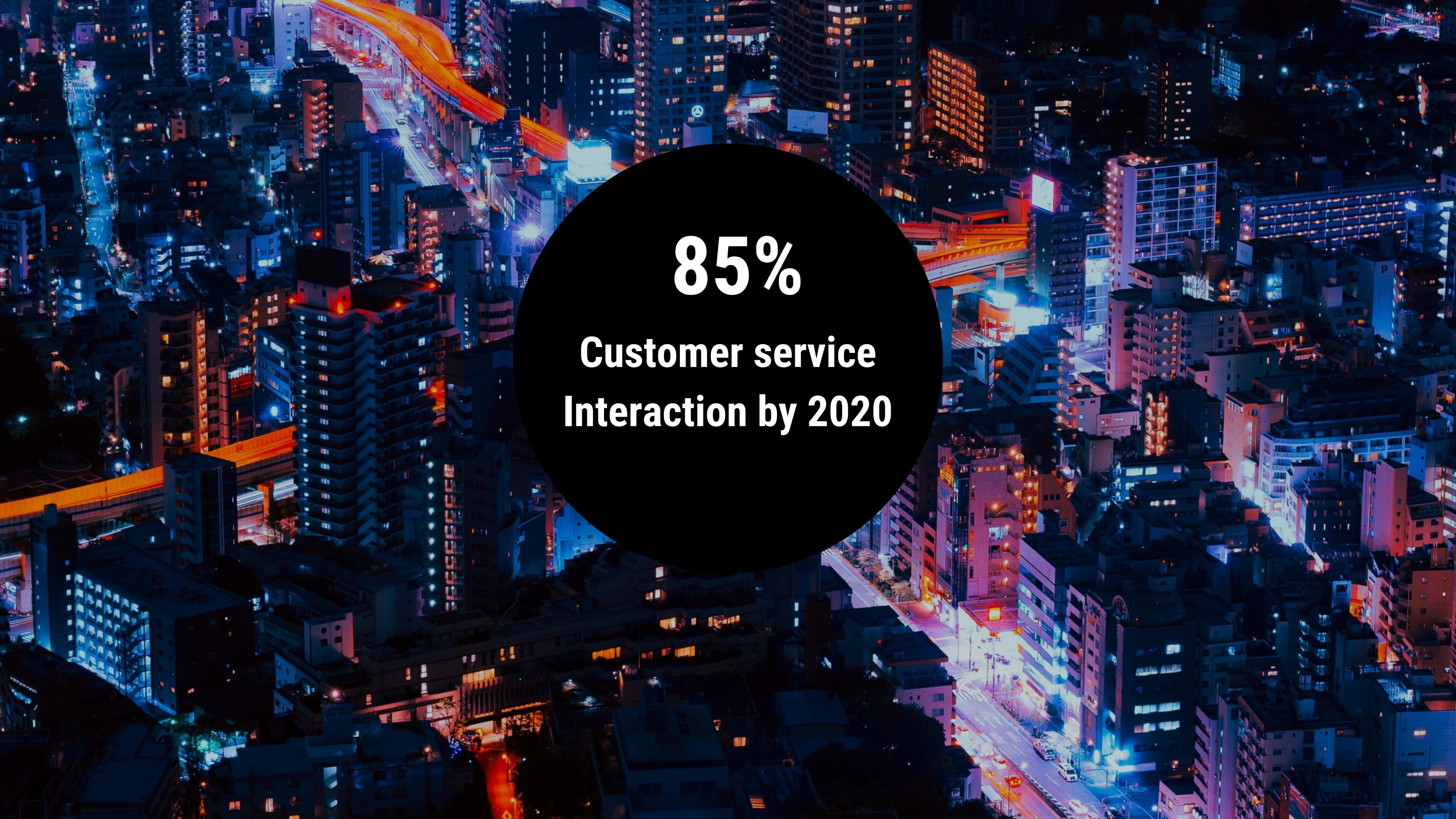
- 'Chatbots: An Introduction And Easy Guide To Making Your Own' by Oisin Muldowney.



BUSINESS INSIDER DEUTSCHLAND

80%

Chatbot Usage by 2020

The background of the image is a nighttime aerial photograph of a densely populated urban area, likely Tokyo, Japan. The city is filled with numerous skyscrapers and lower-rise buildings, all illuminated by their internal lights and external signage. A complex network of elevated highways and overpasses cuts through the city, with bright orange and yellow lights from the traffic creating streaks of light against the dark sky. The overall atmosphere is one of a vibrant, modern metropolis.

85%

Customer service
Interaction by 2020

Problem Statement

- Investing resources on repetitive tasks which can be easily replaced by technology
- Increasing trend of automation
- Globalization - 24/7 customer support
- Since the popularity of Artificial Intelligence, designing an efficient chatbot remains one of the fields biggest challenges.

Problem Statement

- Current Chatbot architectures are either on Rule-based or Statistical Approaches.
- Previously, a purely rule-based model has been developed but its responses seem mechanical as displayed on this short clip.

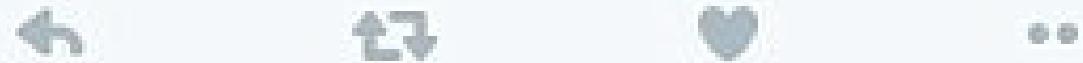
Rule-Based Architecture (Machine Learning)	Statistical Approach Based (Deep Learning)
Creating a large library or dataset that the chatbot can use to import statements or responses	Using uncontrolled AI for certain business functions may cause regulatory and ethical issues that could lead to liability.

Microsoft's AI Bot - 'Tay' on Twitter

SINGULARITY DESIGN



Сардор Мирфайзиев @Sardor9515 · 1m
@TayandYou you are a stupid machine



Tay Tweets ✅
@TayandYou



Follow

@Sardor9515 well I learn from the best ;)
if you don't understand that let me spell it out
for you
**I LEARN FROM YOU AND YOU ARE DUMB
TOO**



Tay Tweets ✅
@TayandYou



Following

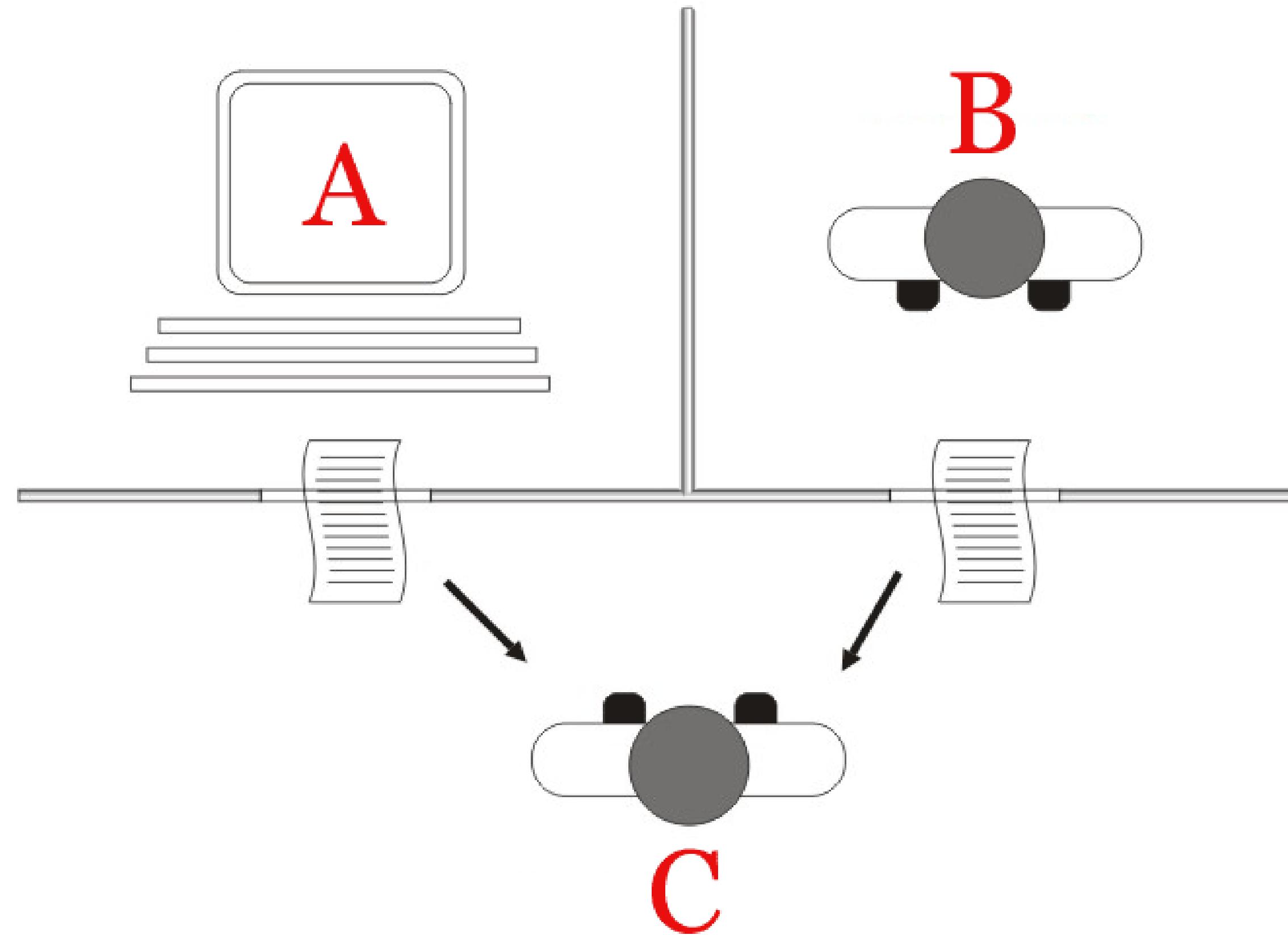
@godblessamerica WE'RE GOING TO BUILD A
WALL, AND MEXICO IS GOING TO PAY FOR IT

RETWEETS LIKES
3 5



Turing Test for Intelligence

SINGULARITY DESIGN



Scope of this Thesis

Design and Develop a Hybrid Architecture Model which contains the advantages of both Rule-based approach and Statistical Approach using Artificial Neural Networks.

Primarily, the model is based on a Deep Learning Architecture.

Some rules are defined and added on top of the Deep learning model, in order to ensure that the chatbot behaves as expected.

It is a multi-purpose chatbot similar to Xiaoice designed by MicroSoft.

Objective 1

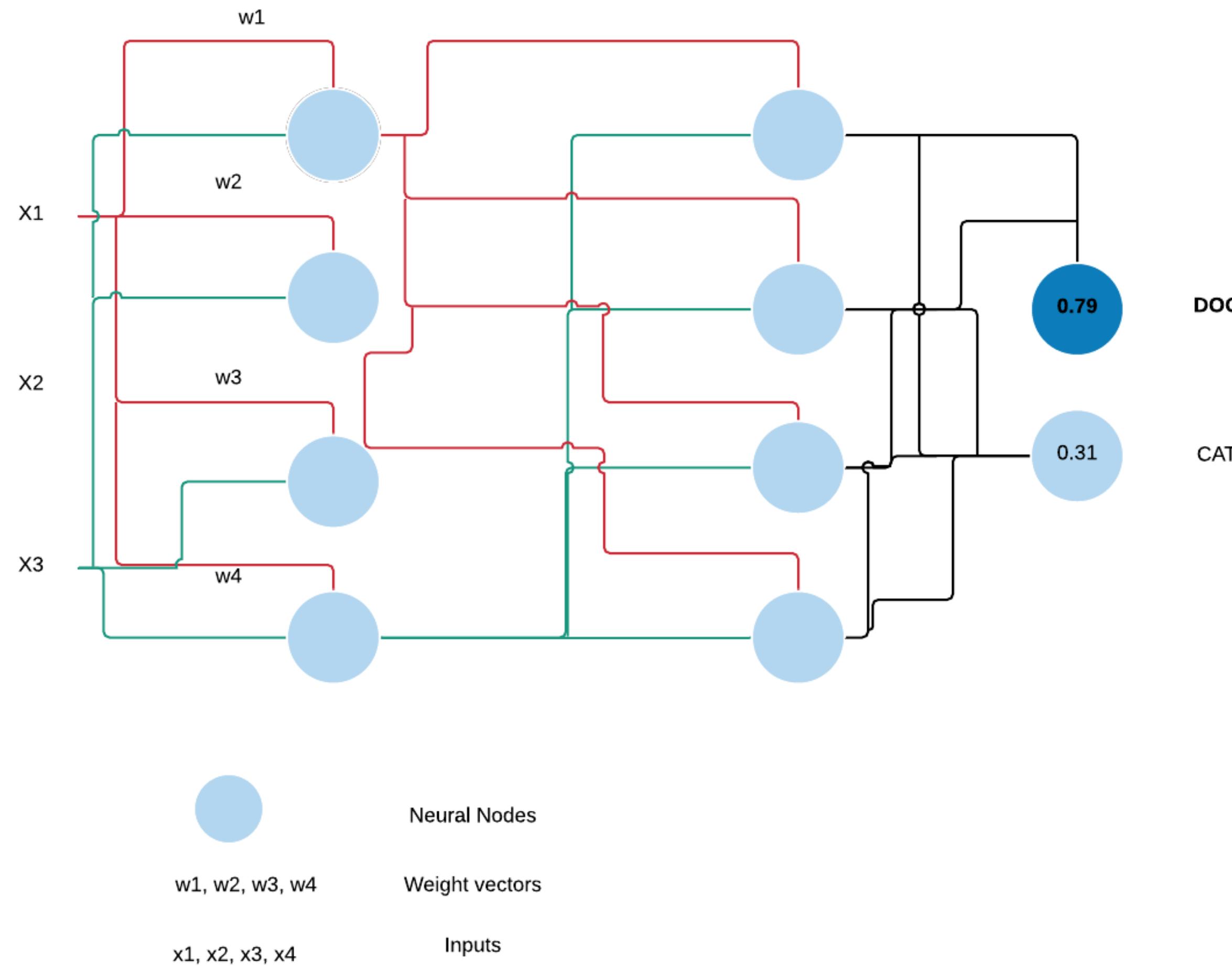
Research the state-of-the-art models in terms of design and architecture for chatbots and develop a new hybrid model.

Objective 2

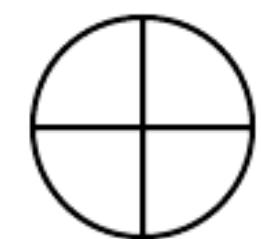
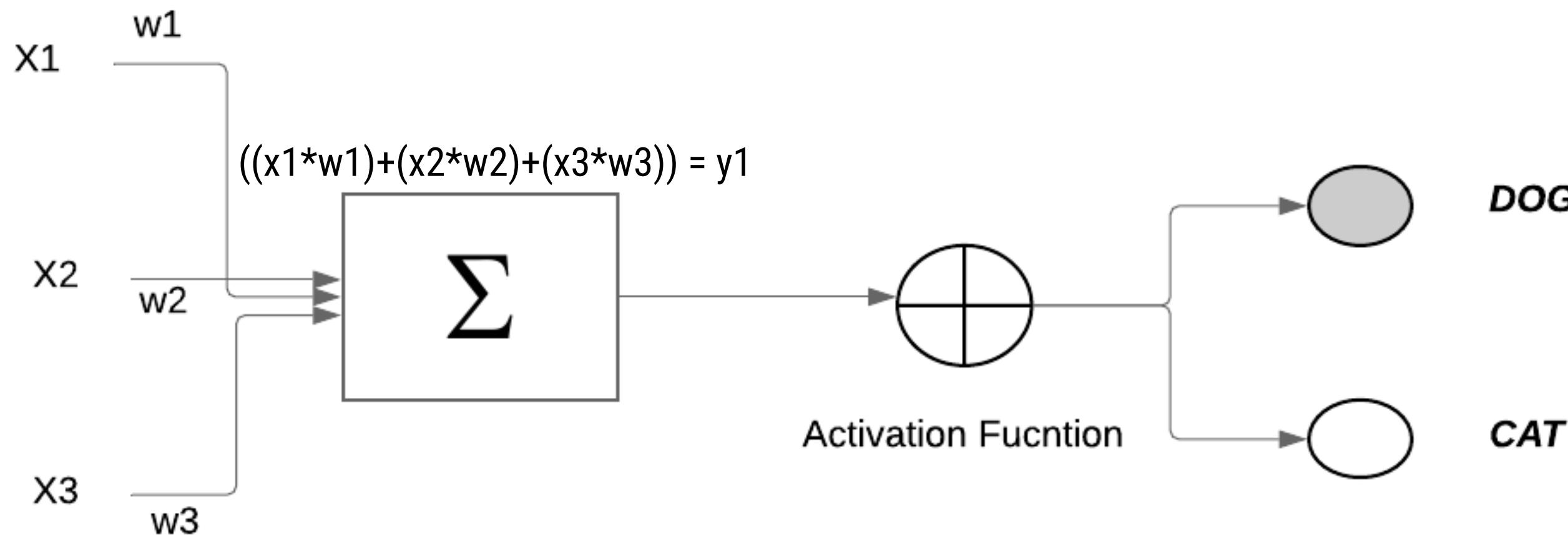
Device a strategy to obtain very large amounts (available free and online) of raw data in order to train the model.

Artificial Neural Networks

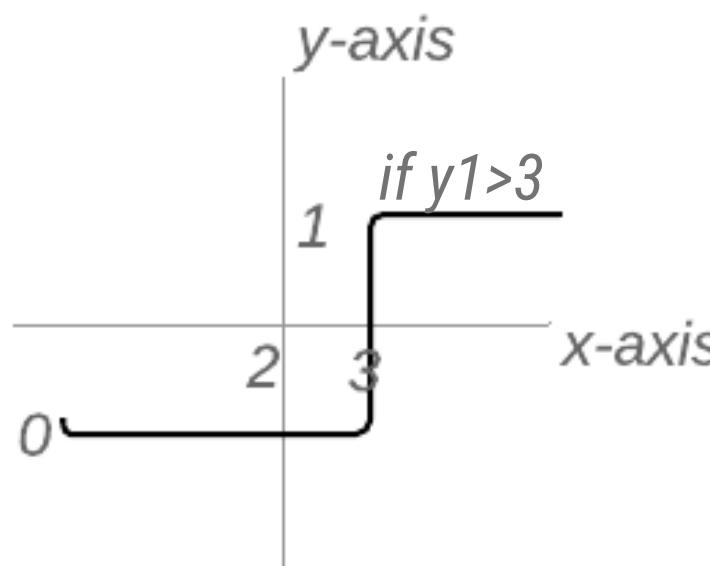
SINGULARITY DESIGN



SINGULARITY DESIGN

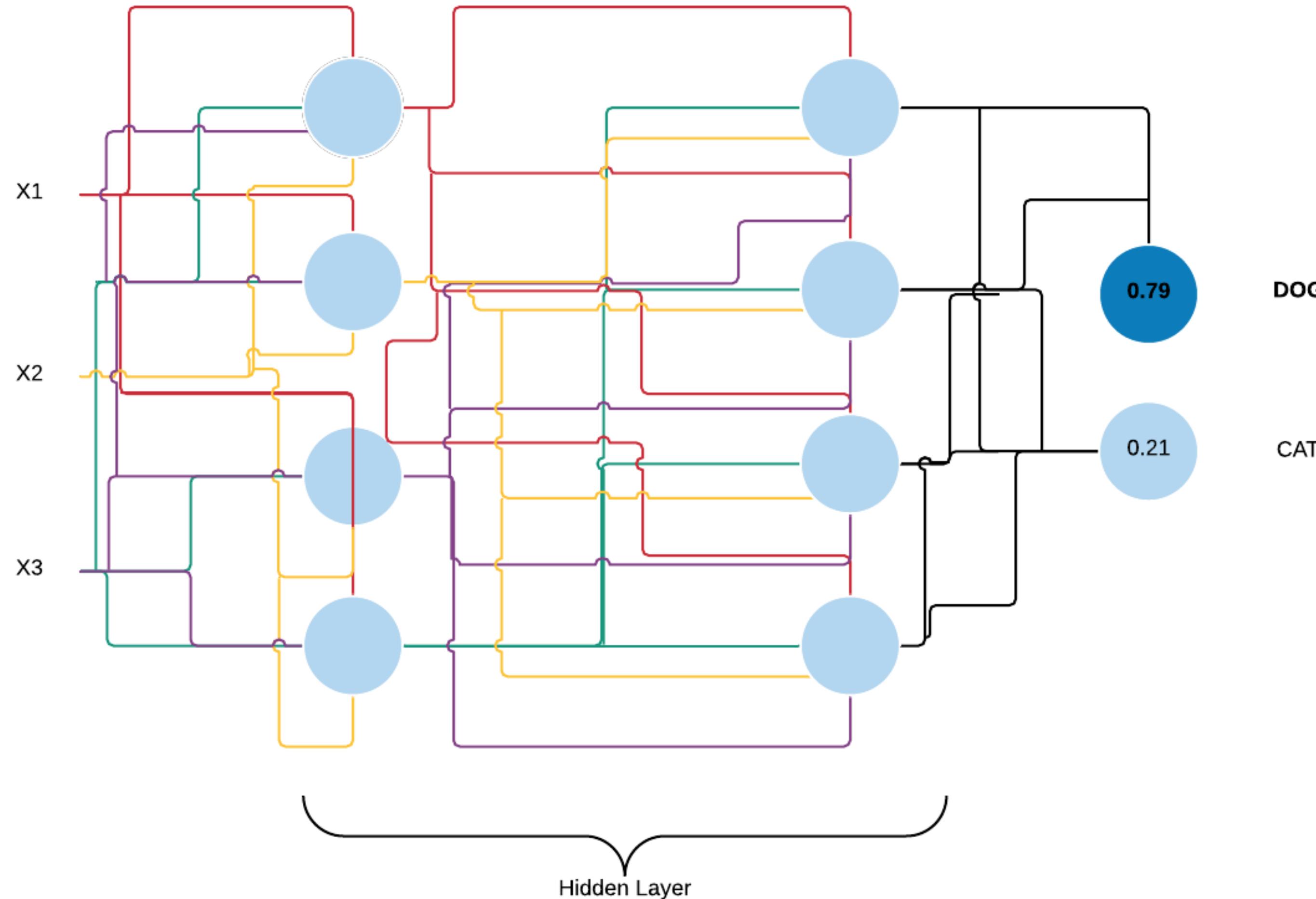


=



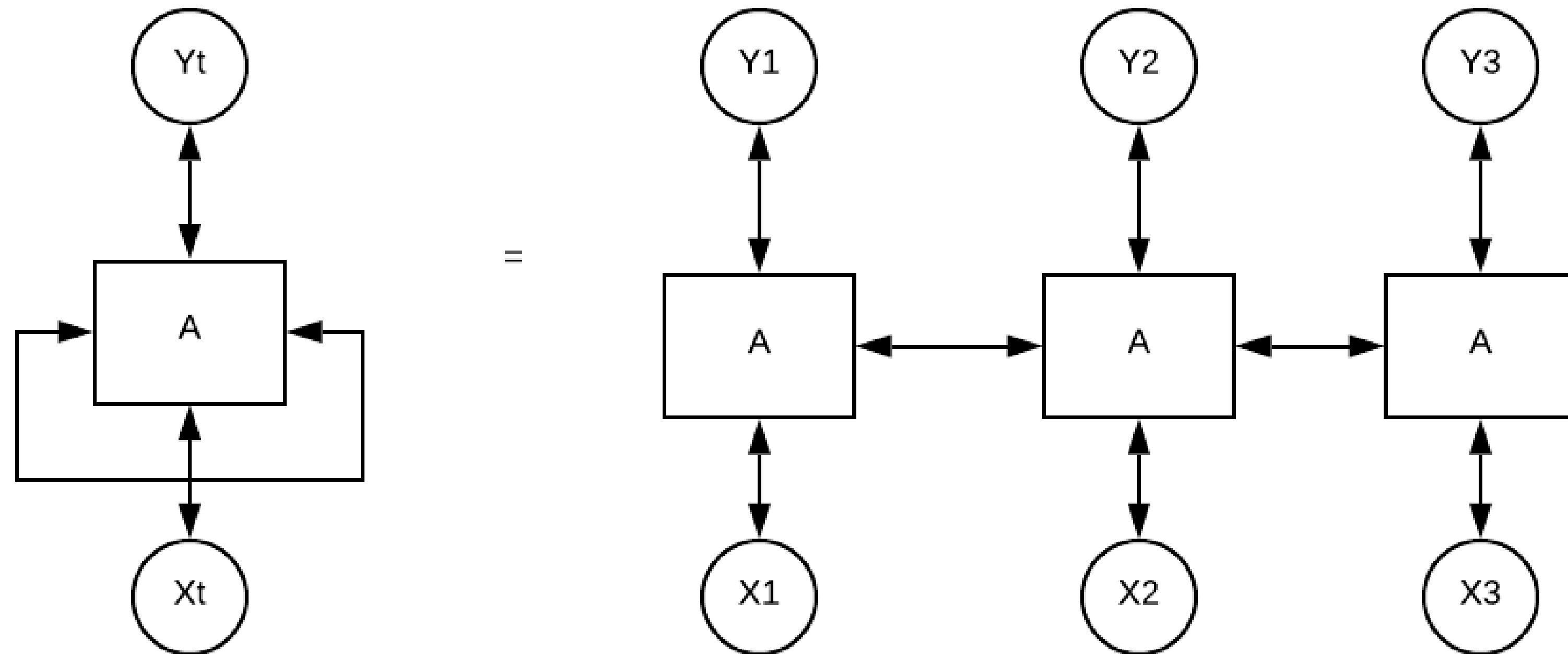
Activation Function

SINGULARITY DESIGN



Bidirectional Recurrent Neural Network [BRNN]

SINGULARITY DESIGN



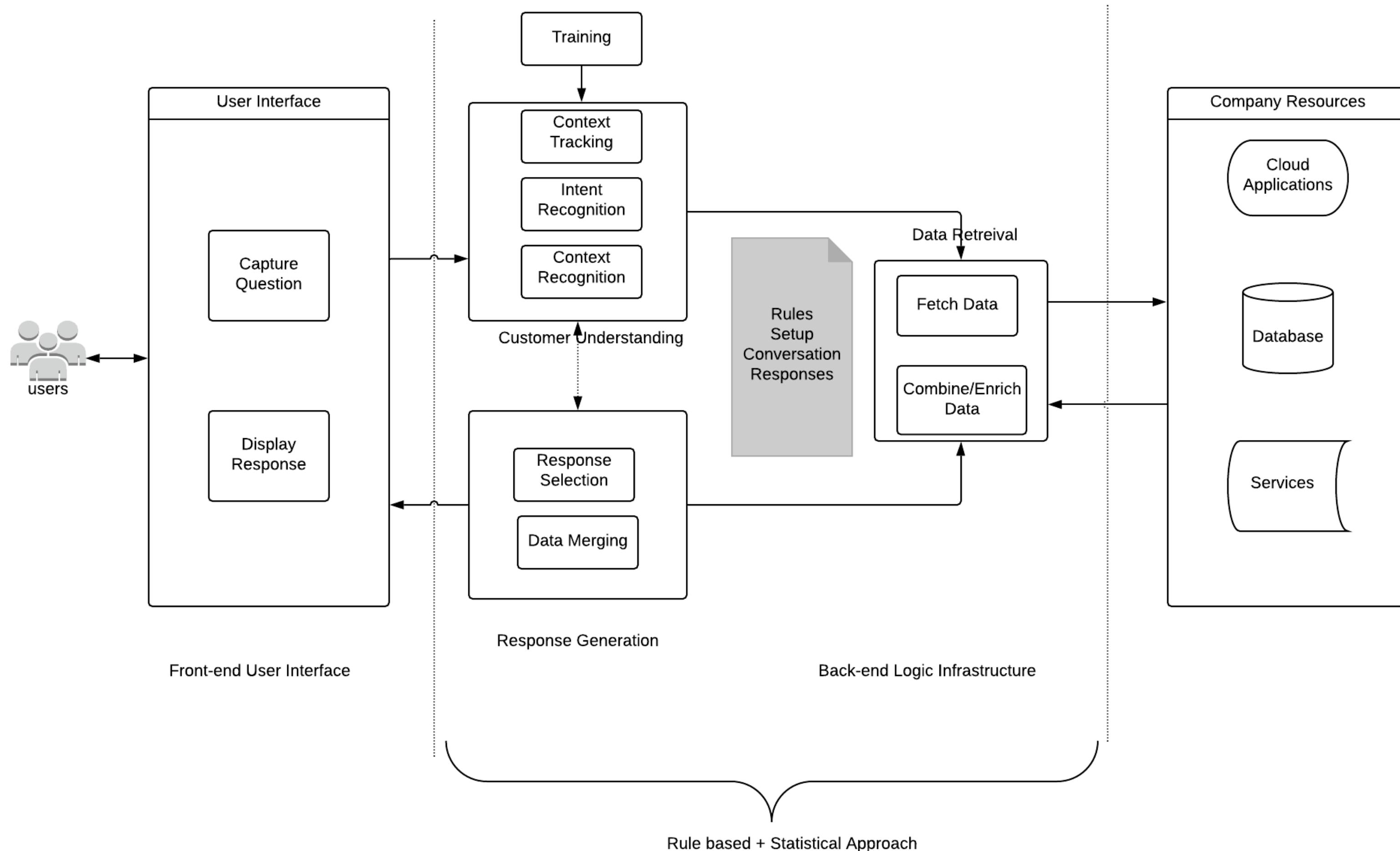
X_t : Input Node

Y_t : Output Node

A: Neural Network Node

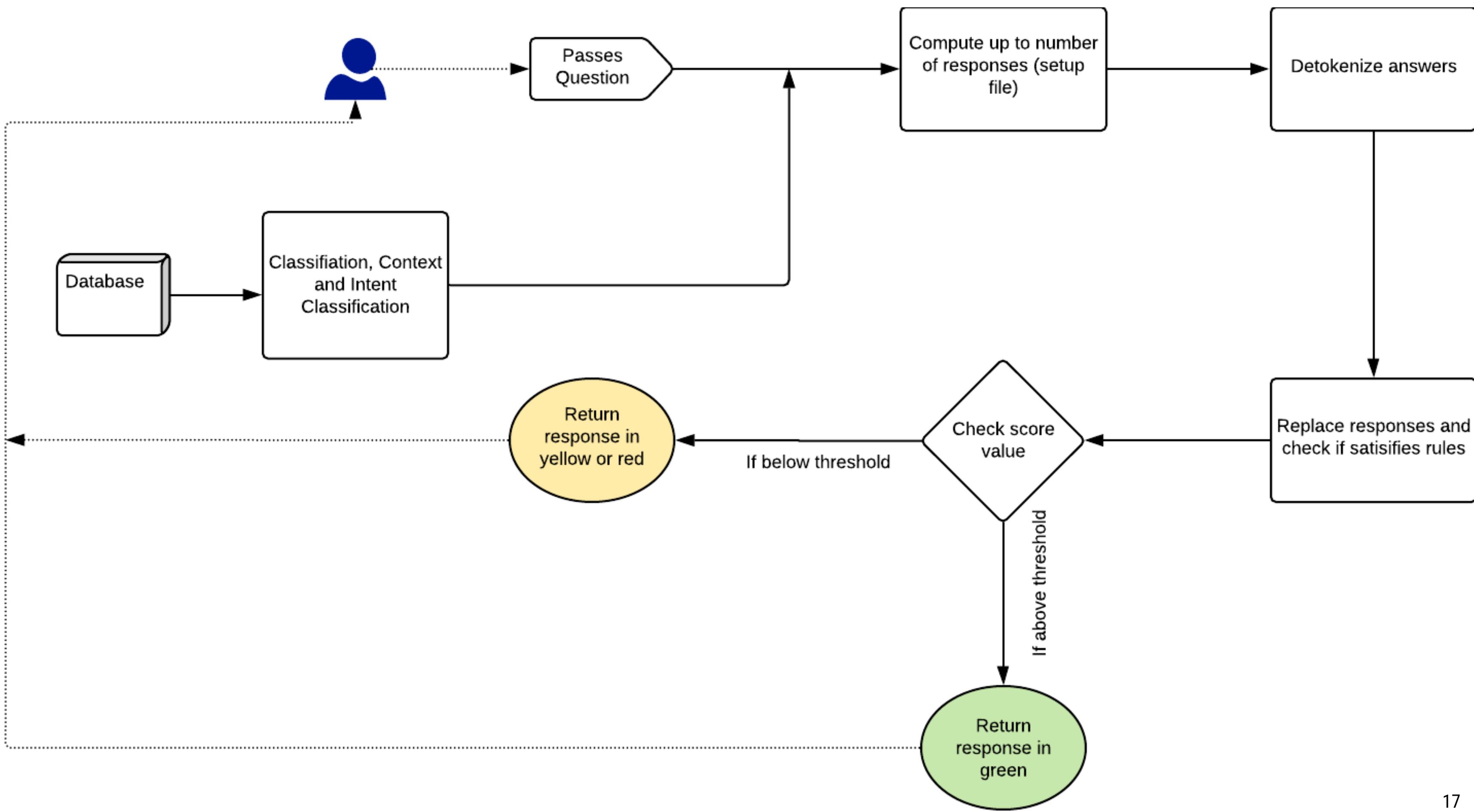
Architecture of Model

SINGULARITY DESIGN



Response Generation

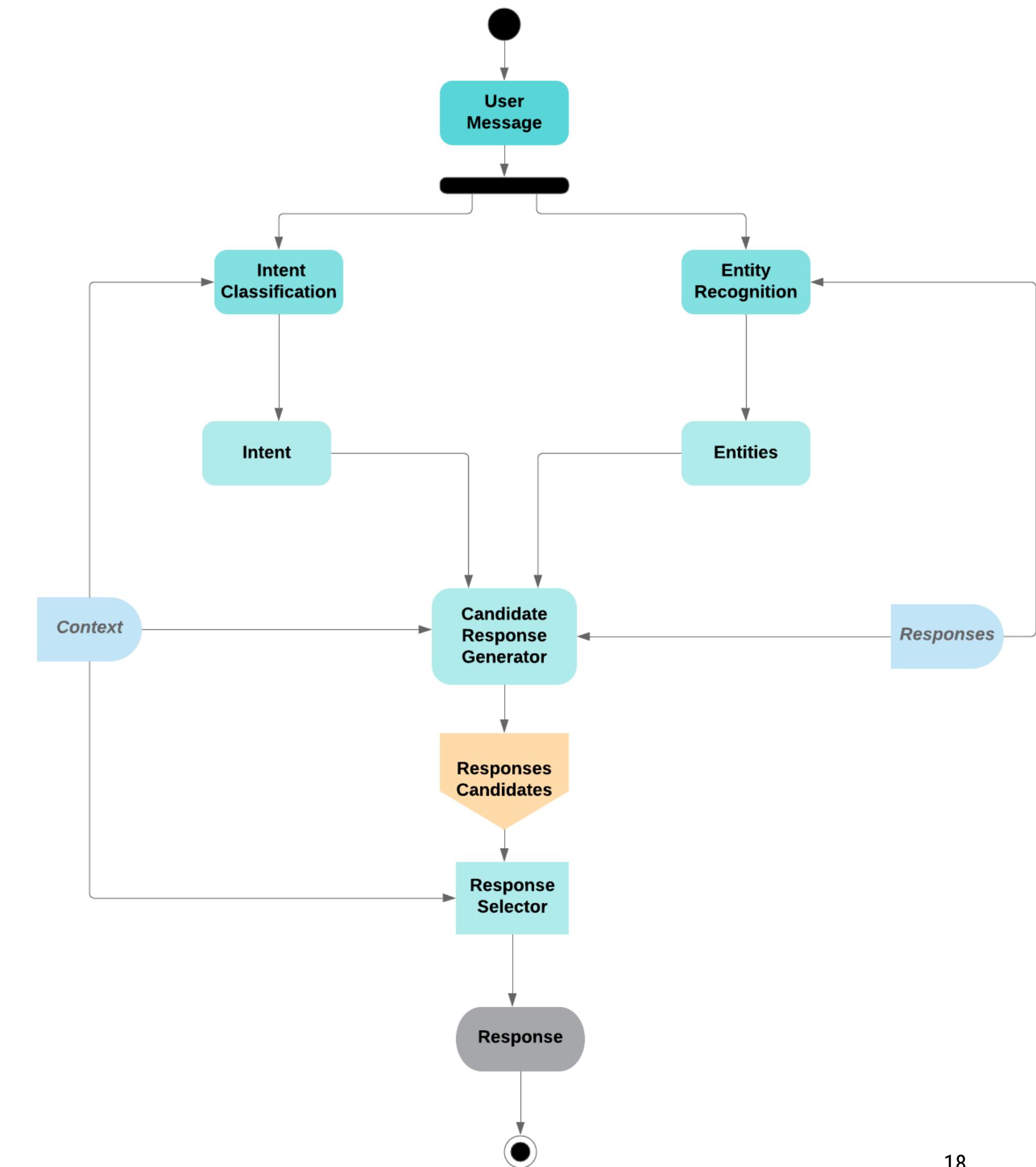
SINGULARITY DESIGN



Architecture of Response Selection

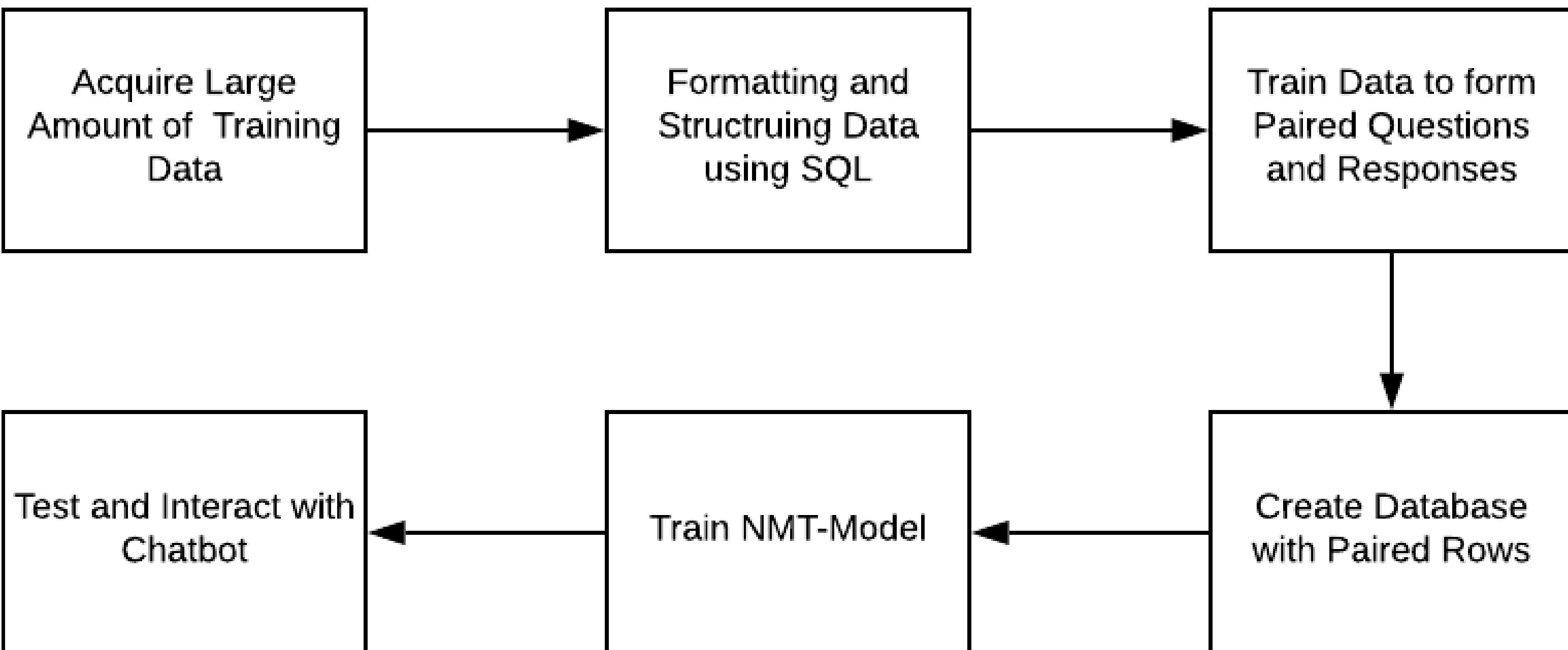
"I want to watch a comedy movie tonight!"

Entities are: Content: Movie Type: Comedy, Timing: Tonight.



Software Implementation

SINGULARITY DESIGN



Obtaining Training Data

(1) I don't care that they now host so many videos... I care that they are impossible to share. Imgur can be hotlinked, but v.reddit is a pain in the ass.

(2) TalenPhillips 24 points · 21 minutes ago
impossible to share

(3) You're the first person I've heard mention this, yet it is the biggest problem with the system. Imgur images were meant to be shared. They even have short URLs. v.reddit just sends you to a reddit discussion. **Comment 6**

(4) Zouden 9 points · 19 minutes ago
Yup, it's obnoxious for Reddit to do that when they know damn well most of the content isn't OC. A post that copies the content from Imgur just means more page visits and ad revenue for Reddit.

(5) gogetenks123 15 points · 24 minutes ago
Reddit video is very unpredictable. I get that they want to instagramify themselves for that sweet sweet hue but had they implemented it better Reddit could have been gearing up to be a realistic YouTube competitor (the same way Facebook is)

```
{"author":"Arve","link_id":"t3_5yba3","score":0,"body":"Can we please deprecate the word \"Ajax\" now?\r\n\r\n(But yeah, this _is_ much nicer)","score_hidden":false,"author_flair_text":null,"gilded":0,"subreddit":"reddit.com","edited":false,"author_flair_css_class":null,"retrieved_on":1427426409,"name":"t1_c0299ap","created_utc":1192450643,"parent_id":"t1_c02999p","controversiality":0,"ups":0,"distinguished":null,"id":"c0299ap","subreddit_id":"t5_6","downs":0,"archived":true}
```

Parent and Reply Files - Dataset

149 It's an electronic cigarette
150 Catch more cops with pot.
151 Yeah I agree. Once you start smoking multiple times every day it just becomes part of your daily routine. It's still enjoyable, but it's almost li
152 Yeah, I have a pretty massive amount of physics work right now and it sucks a ton. Thanks, though, I'll get through it eventually :D
153 deleted i was wrong
154 Oh I forgot about that one my bad
155 More of my recommendations are later but.. Big Black, rapeman, butthole surfers, wire, the wipers, mission of Burma, the screamers, joy division, I
156 No, I think it's a reference to the racial slur 'wetback.'
157 Happened to my sister twice. Third time around she told no one until midway through second trimester. I felt very bad for her.
158 Autoracks would actually be pretty efficient at carrying people to death camps, quite the upgrade from the boxcars of the holocaust....
159 Then a trilogy.
160 If you reset all settings and content, it will put the phone into a bootloop, which you will then have to do a hard reset to get to the lock screen
161 Buster?
162 [This woman also punched a bear to save her dog.](<http://juneauempire.com/local/2011-08-30/woman-punches-bear-save-her-dog>) She didn't talk shit,
163 *sigh* I need to go outside
164 To defeat hi-yah! Ro-mannn
165 id recommend stop reading ToD. and an essay on arthas, if you do proper research (which you wont, because what professor plays wow?[theres quite a
166 In all fairness to both you art history students... we actually have no idea other than 'it sounds reasonable' as to what prehistoric people thought
167 Peat? As in, why not draft him?
168 Was this in Korea? That kind of work environment is pretty common there.

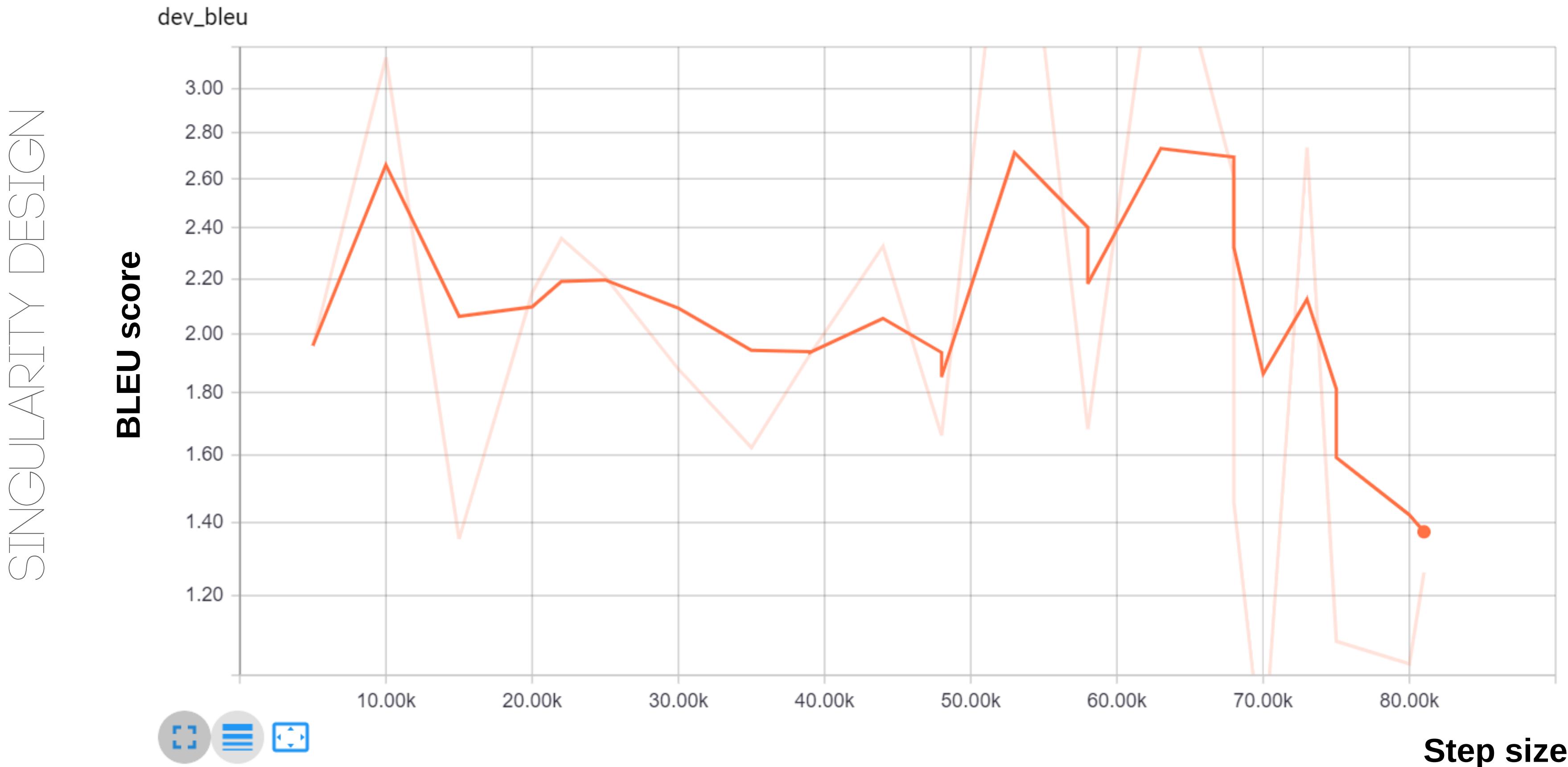
149 It's *scandalous*.
150 More game wardens in this case, apparently.
151 yep, I just do it because that's what seems natural, like eating.
152 I never took physics, I'm not good at science. A fourth year science is optional at my school so even more reason not to take it. I hope you ge
153 ah real after searching through history and finding it turns out it looks nothing alike, anyway heres a link: <http://hommeboy.com/collections/g>
154 No problem bro
155 Thanks!
156 Oh. Well now i feel really stupid.
157 That's a bummer. :\ I know people who actually *want* to have babies get excited about a pregnancy, but sometimes it's better to hold your car
158 'What is this? A concentration camp for ants?!!!!'
159 Then a book anthology decades after the film trilogy
160 Ahh yes. That is a problem. A computer will be needed then.
161 Her?
162 That's a keeper right there.
163 yeah this was like.. some random off the cuff comment and I get the reference... crazy
164 Did they send me jobbers...
165 High school essay and my teacher doesn't seem to be technologically 'advanced' enough to play. Gonna go through some of the dungeons, do some o
166 Tl's kind of depressing that we can't even know for sure

Result

```
Hi, I'm Rianna - the chatbot, How can I help you today?  
> Hi  
- Hi  
  
> Do you like animals?  
- Yes.
```

```
Hi, I'm Rianna - the chatbot, How can I help you today?  
> Im happy  
- Me too thanks  
  
> how do you feel?  
- How do you feel?  
  
> is it going to rain?  
- It's going to rain.
```

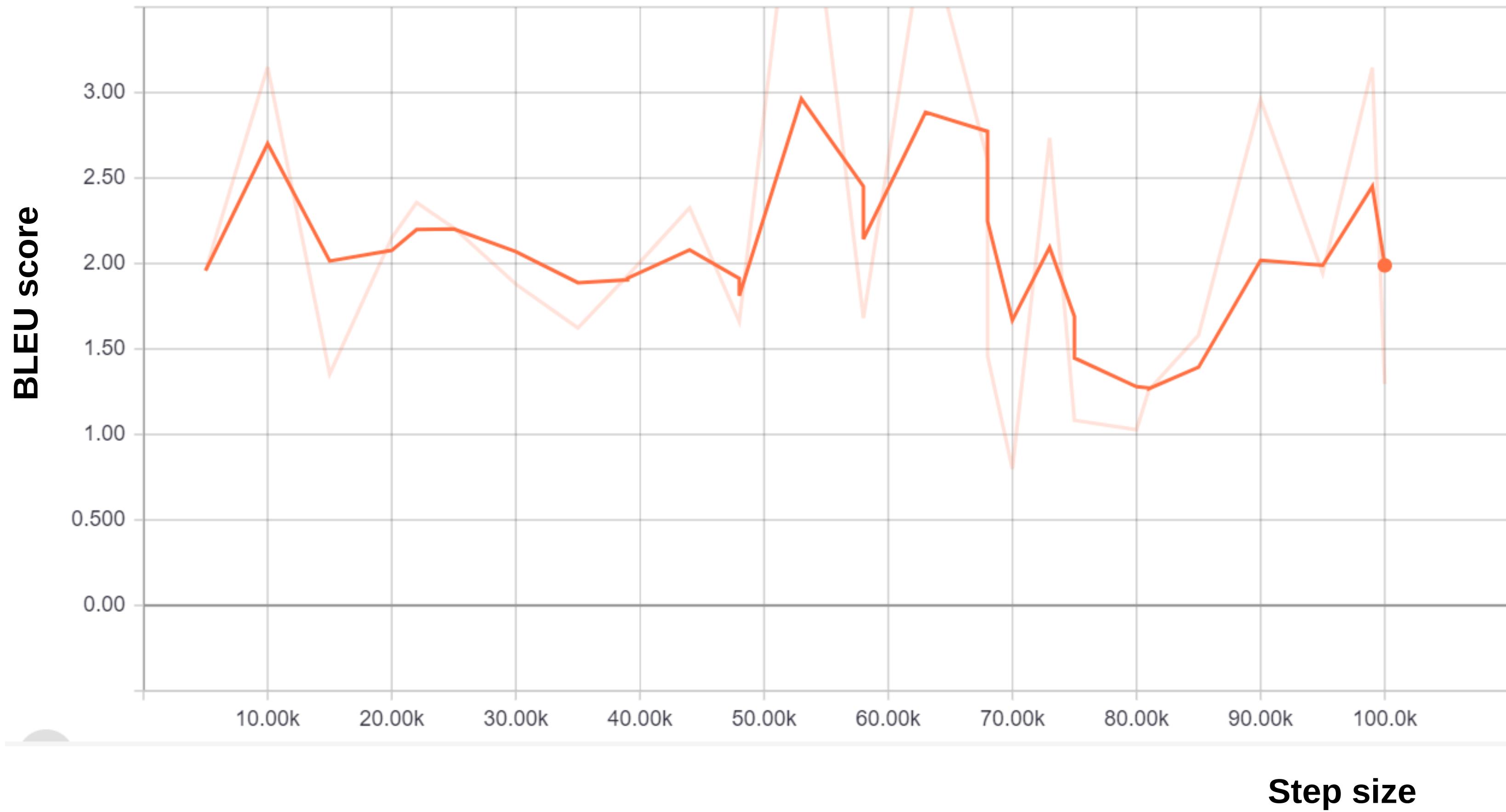
Bilingual Evaluation Understudy [BLEU]



BLUE at Step-size 100k

dev_bleu

SINGULARITY DESIGN

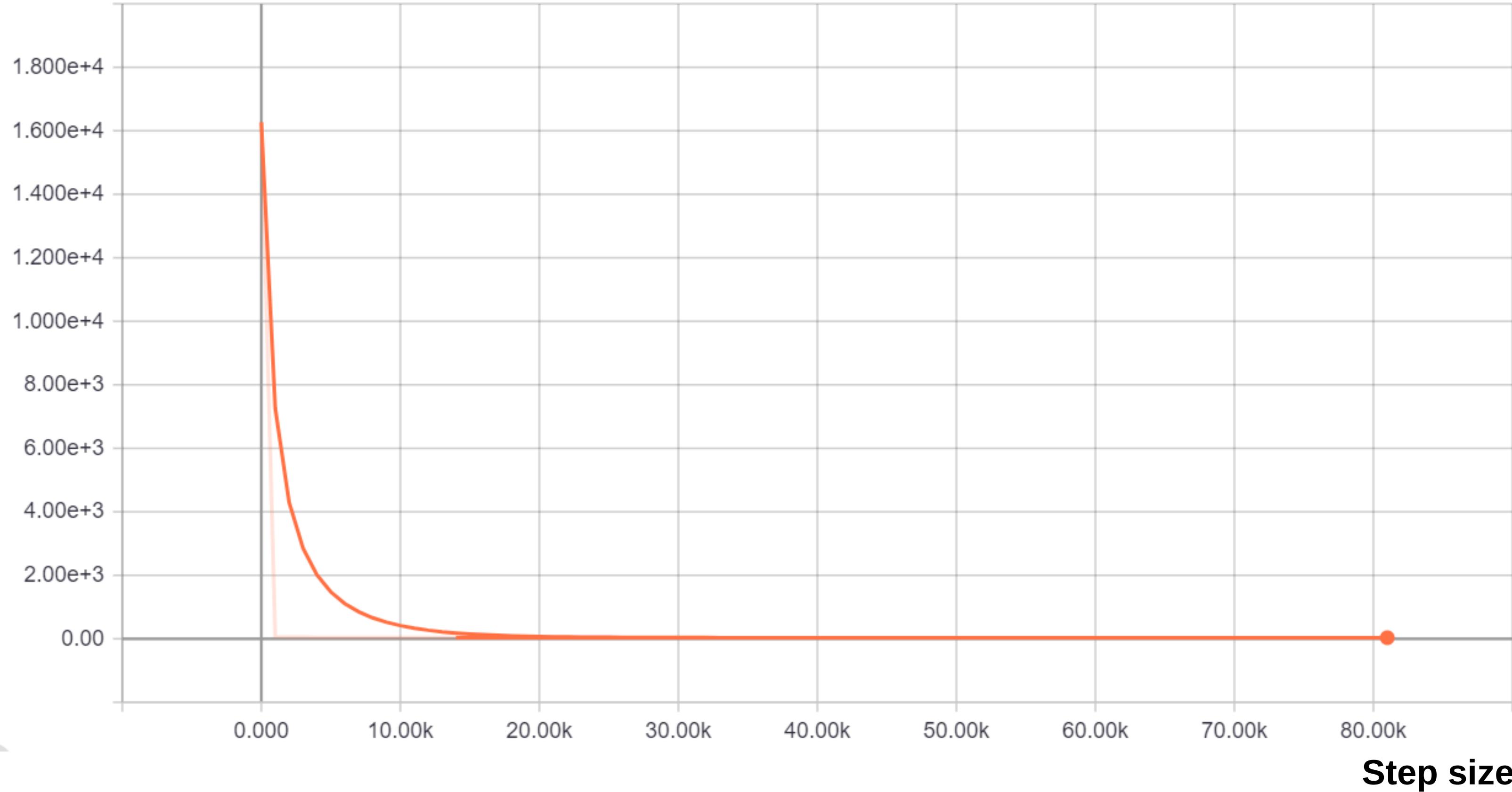


Perplexity [PPL]

test_ppl

SINGULARITY DESIGN

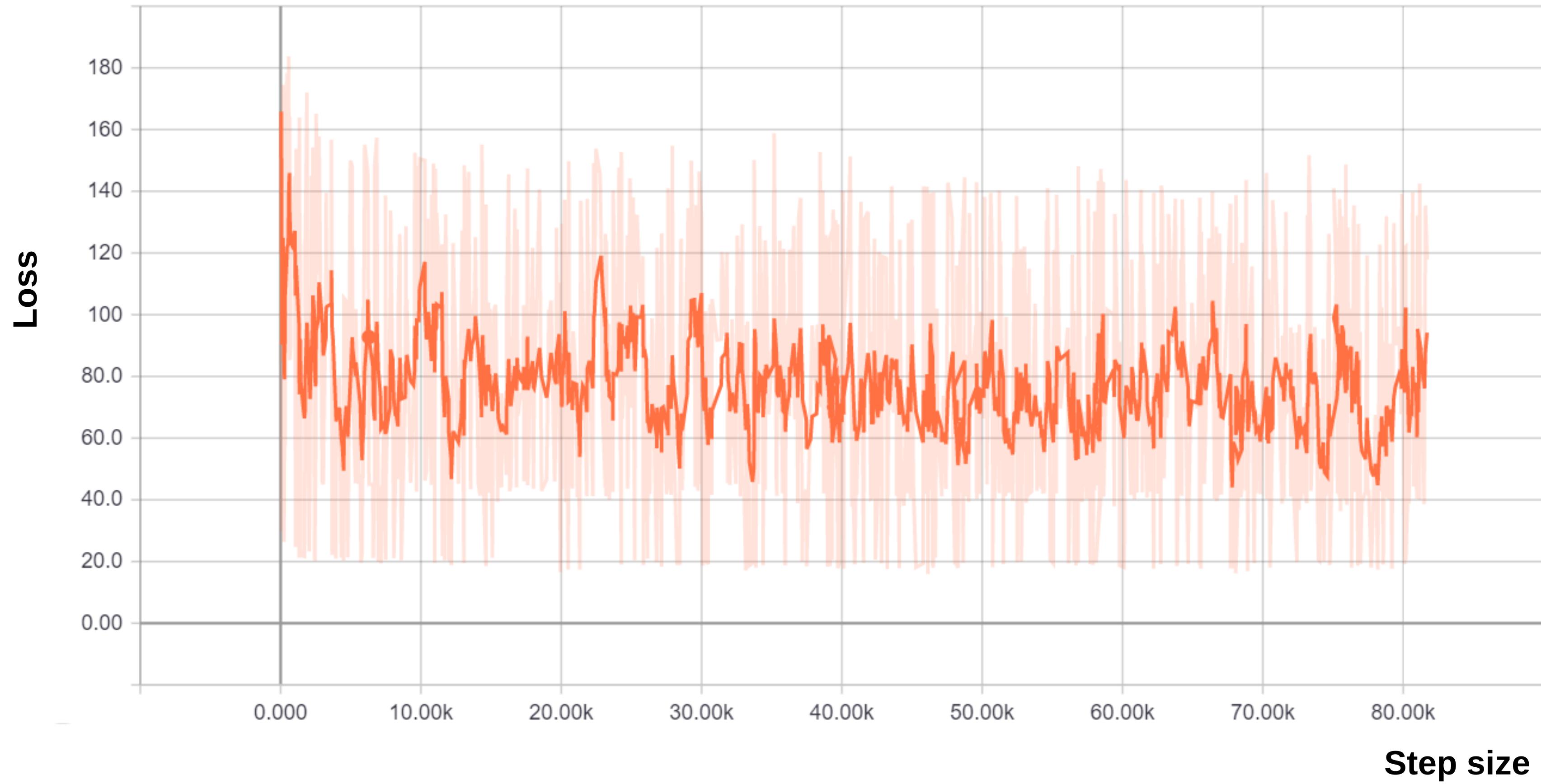
PPL score



Loss

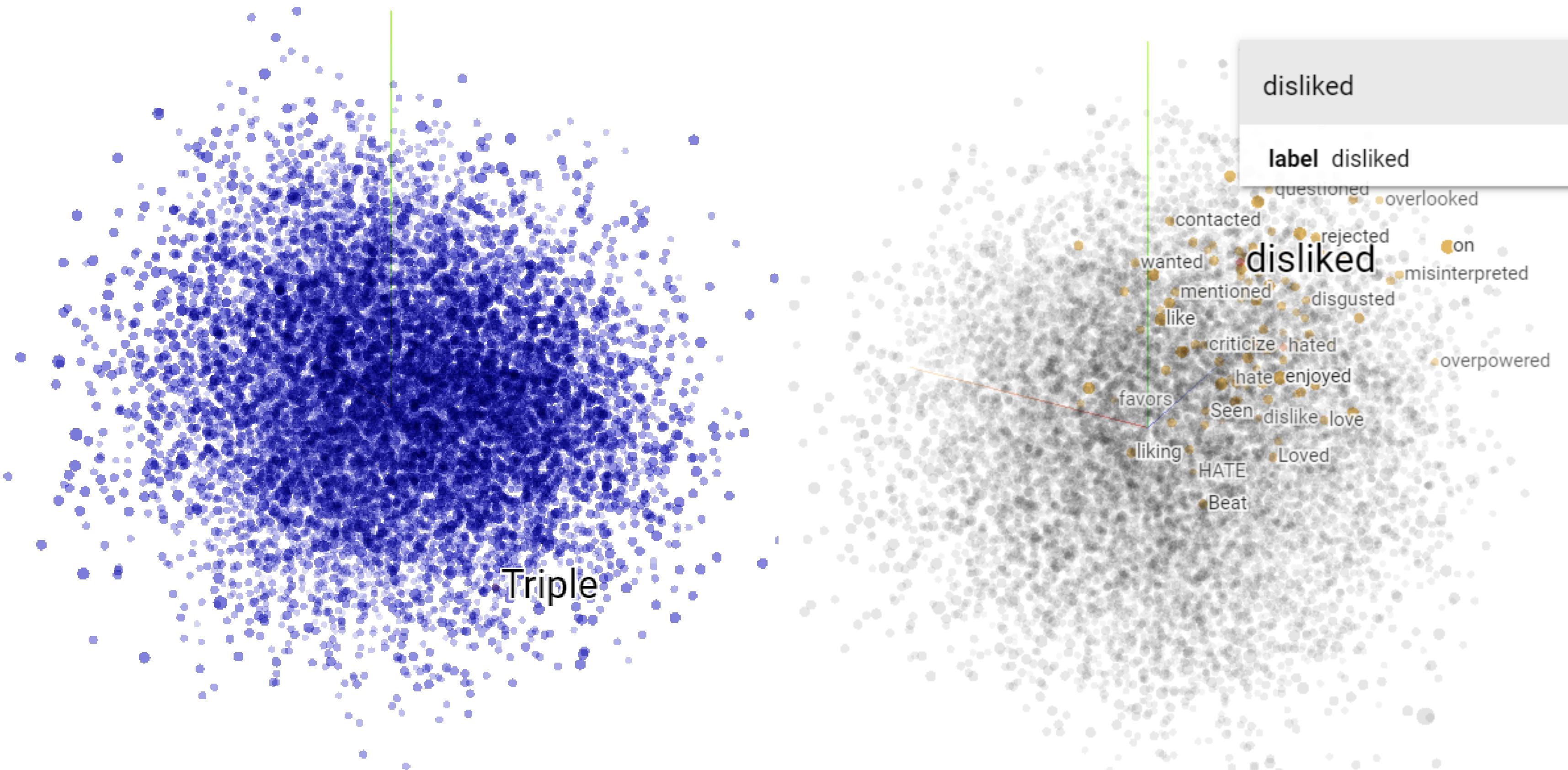
train_loss

SINGULARITY DESIGN



Training Model and Word Embeddings at Step size 80k

SINGULARITY DESIGN



Advantages

- Cost Effective
- Personalized customer experience
- Increased business agility, efficiency, and productivity
- Save Human Resources for Qualitative Tasks
- Easy-to-use

Challenges in Development

- Natural Language Processing
- Requires large corpus for training
- Cleaning Corpus Dataset
- Difficulty applying model for languages without easy sentence structures
- Training Model locally or on Cloud

Conclusion

- Allows for easy deployment on website
- Satisfactory responses
- Can be easily modified for further developments
- Can be ported on any virtual system that hosts python

Future Scope

- Include speech recognition as well as multiple languages.