



Conversation Topic Classification at Talkdesk

Pedro Verruma
Engineering Manager

powered by
talkdesk



Who am I?

Pedro Verruma

pedro.verruma@talkdesk.com

Engineer Manager at TDX

talkdesk.com/tdx

NLP Background

10 years of experience in AI and ML



Agenda

Meet the Company

Talkdesk, TDX

Conversation Topics?

Context and Motivation

Sentences

Classification

Moments

Classification





MEET THE COMPANY

TALKDESK & TDX



Company

talkdesk

Empower companies to continuously improve customer experience.

Mission

1.800+

Customers

30.000+

Users

800+

People



Innovation



TDX is an Innovation Lab to fast-track key technology initiatives and accelerate technological responses to business issues.

Mission

Coimbra | Aveiro

Location

1/2019

Founded

80+

People





CONVERSATION TOPICS?

CONTEXT AND MOTIVATION



Context

Talkdesk handles a **lot** of calls.

Calls are transformed into text, called **transcripts**.
This allow for easier manipulation and information extraction.

What are **conversation topics**?



Call Transcript



Motivation



Why should we get topics from conversations?

Human agents already do this but is prone to **errors** and **time consuming**.

We want to automate it, allowing extraction of **metrics**:

- Number of calls about a specific topic.

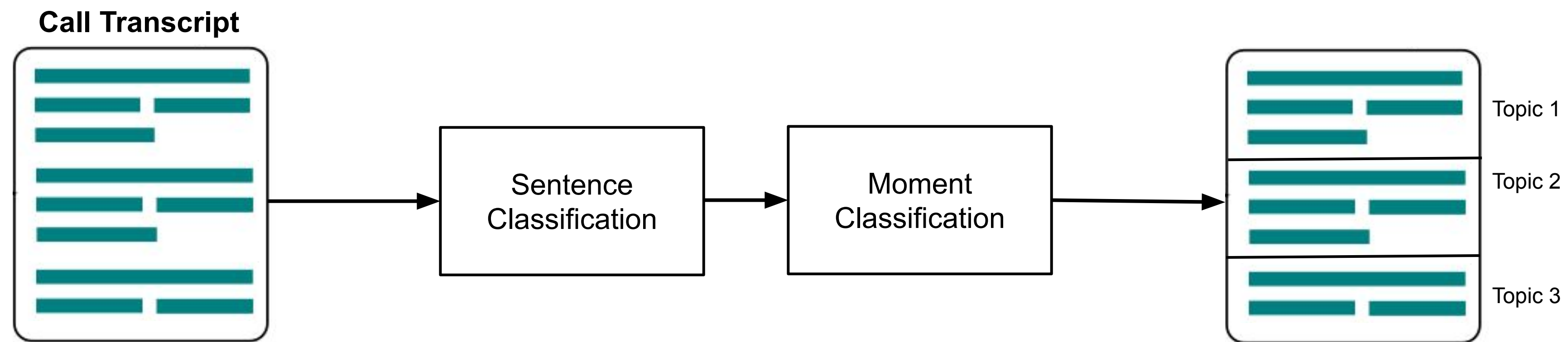
- Number of calls with multiple topics.



How to handle it?



Sentence and Moment Classification





SENTENCES

CLASSIFICATION



What do we have?

Call **transcriptions**.

Separated by **Agent** and **Customer**.

Splitted into **sentences**.

Each sentence is **classified** individually.



What do we do?

For each **sentence** we classify it with an **intent**:

GREETING
FAREWELL

INFORMATION_REQUEST
INFORMATION_RELAY
INFORMATION_ACKNOWLEDGEMENT

MOMENT_SHIFT



Real Scenario



Call Transcript		Intents
Moment 1	Hi!	Greeting
	I am currently planning to come to Cambridge, and I was looking for a relatively inexpensive place to eat in the centre.	Information Relay
	What would you suggest?	Information Request
	The Gardenia is an excellent and cheap place in the center of town, serving Mediterranean cuisine.	Information Relay
	Yes, please book a table for 1 people at 19:15 on monday.	Information Relay
Moment 2	Reference number is : IU63HAEN.	Information Relay
	Is there anything else I can help you with today?	Moment Shift
	I need a place to stay in the same part of town.	Information Relay
	I must have free parking.	Information Relay
	We have two guesthouses and two hotels that meet your needs.	Information Relay
Moment 2	...	





Approach



Dataset



MultiWOZ

Fully-labeled collection of written conversation spanning over multiple domains and topics.

<https://github.com/budzianowski/multiwoz>



MultiWOZ



```
{
  "utterance_text": "Your booking for one ticket is complete",
  "intent": "INFORMATION_RELAY"
},
{
  "utterance_text": "Your reference number is VEG5Q87Q and 75 GBP will be due at the station",
  "intent": "INFORMATION_RELAY"
},
{
  "utterance_text": "I am also looking for an attraction called old schools",
  "intent": "MOMENT_SHIFT"
},
{
  "utterance_text": "Yes, Old Schools is located in the centre area, and has no entrance fee",
  "intent": "INFORMATION_RELAY"
},
{
  "utterance_text": "Anything else I can assist you with?",
  "intent": "MOMENT_SHIFT"
},
}
```



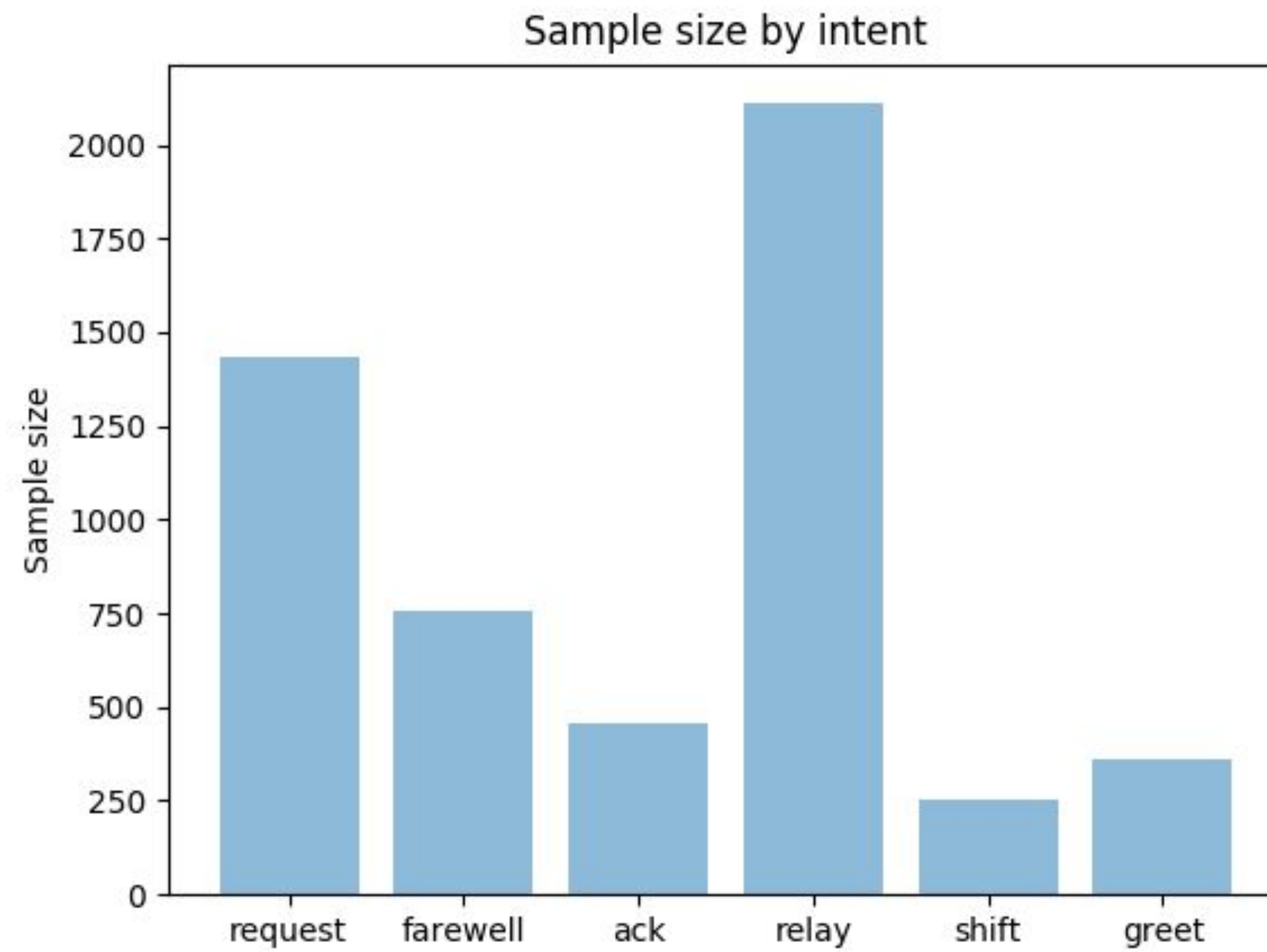
MultiWOZ



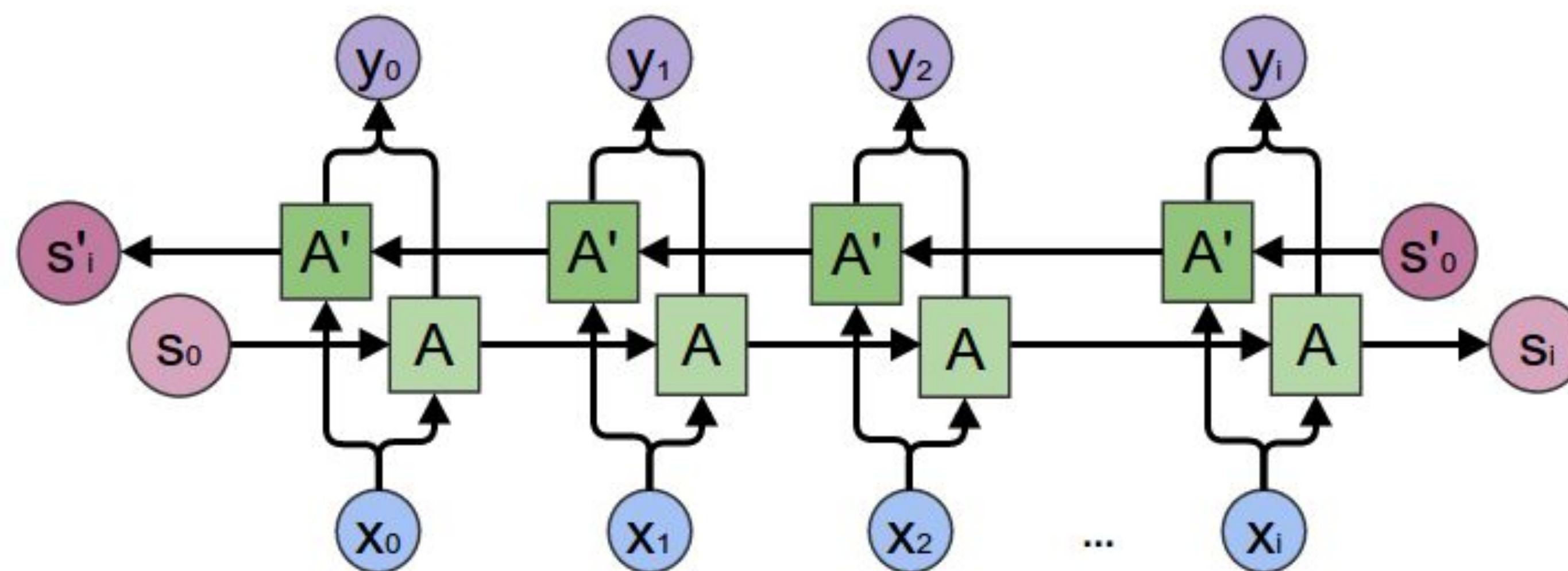
```
{
  "utterance_text": "Can you please help me get in touch with the police?",
  "intent": "INFORMATION_REQUEST"
},
{
  "utterance_text": "Parkside Police Station is in Parkside, Cambridge",
  "intent": "INFORMATION_RELAY"
},
{
  "utterance_text": "Their number is 01223358966",
  "intent": "INFORMATION_RELAY"
},
{
  "utterance_text": "Anything else I can do for you?",
  "intent": "MOMENT_SHIFT"
},
{
  "utterance_text": "Can I please have the postcode as well?",
  "intent": "INFORMATION_REQUEST"
},
{
  "utterance_text": "The postcode for the Parkside Police Station is CB11JG",
  "intent": "INFORMATION_RELAY"
},
}
```



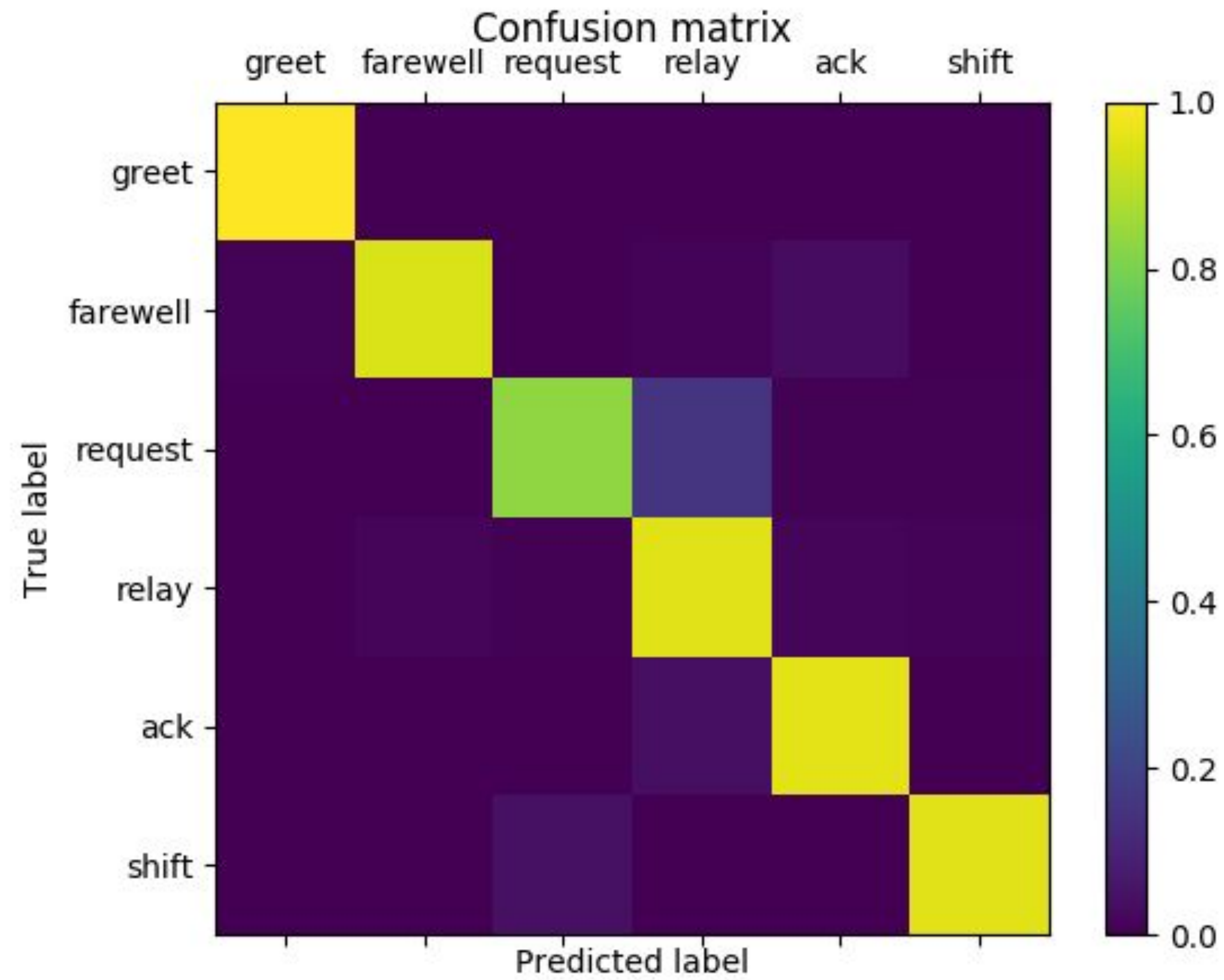
Dataset by intent



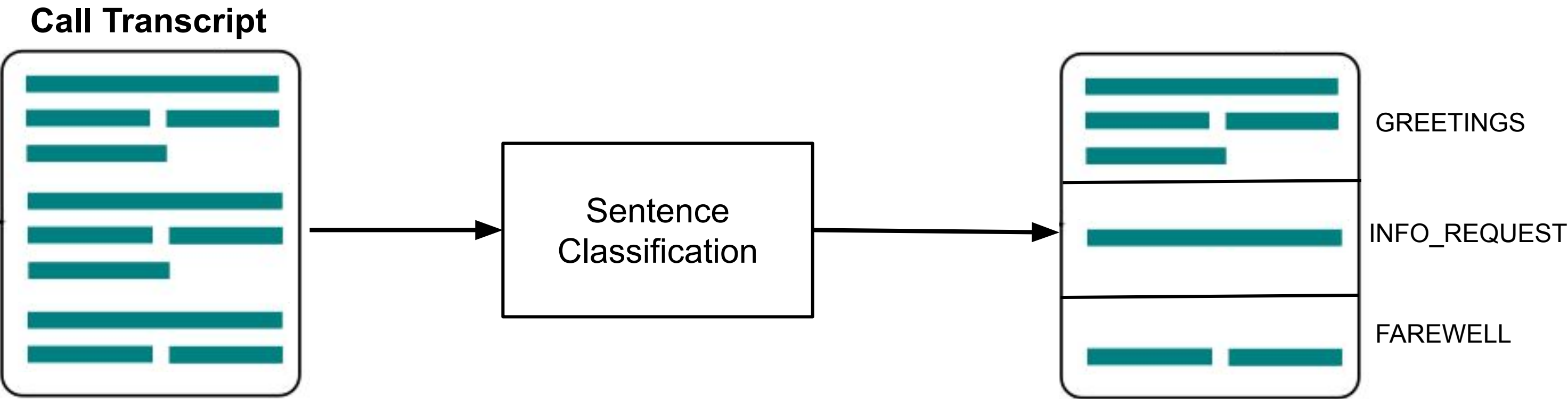
Model



Results



Recap





MOMENTS

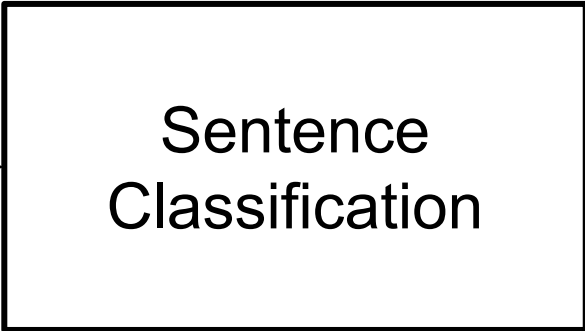
CLASSIFICATION



Context & Goal



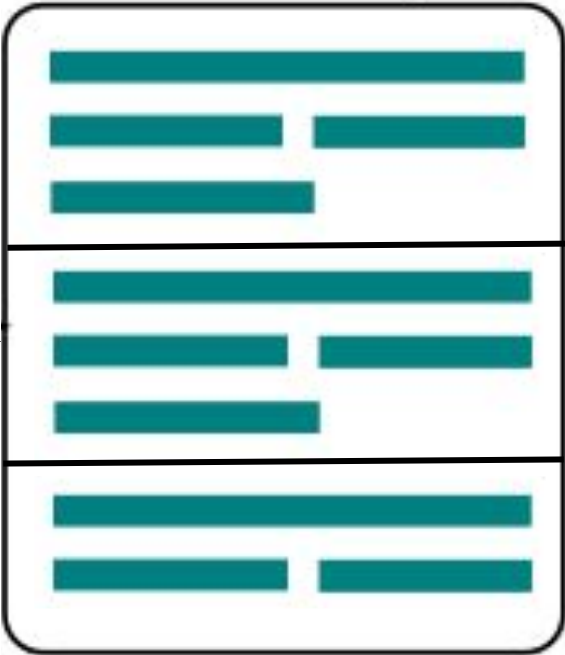
Call Transcript



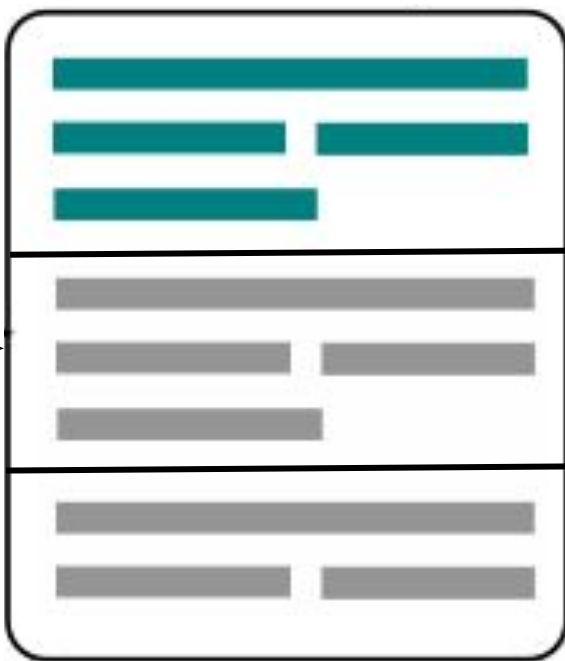
Moment 1

Moment 2

Moment 3



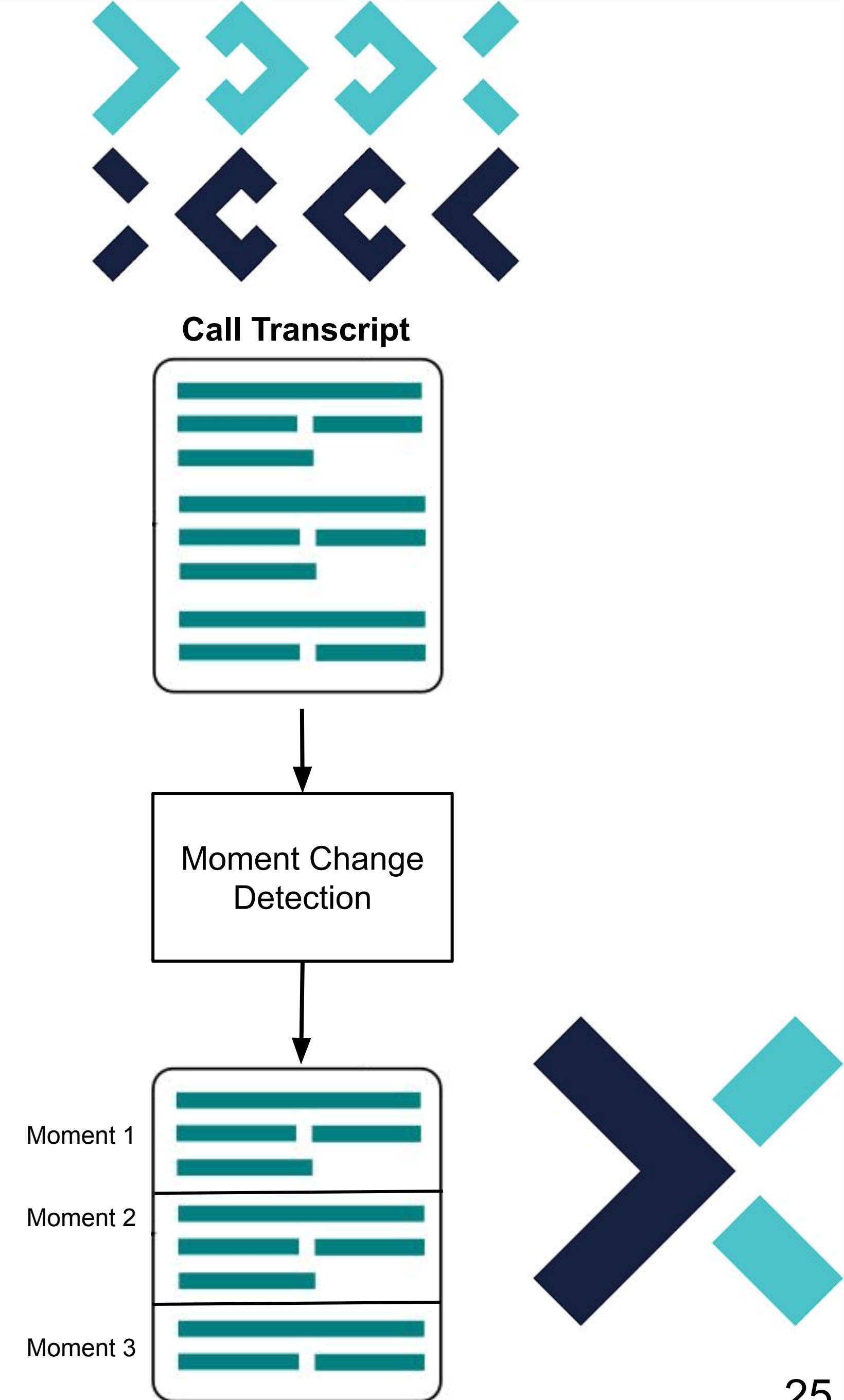
Return order



Moment Change Detection

Uses the **Sentence Classification** to detect the intent of each sentence, and splits the transcripts in **Moment Shift intents**.

Model to detect intents trained a priori.

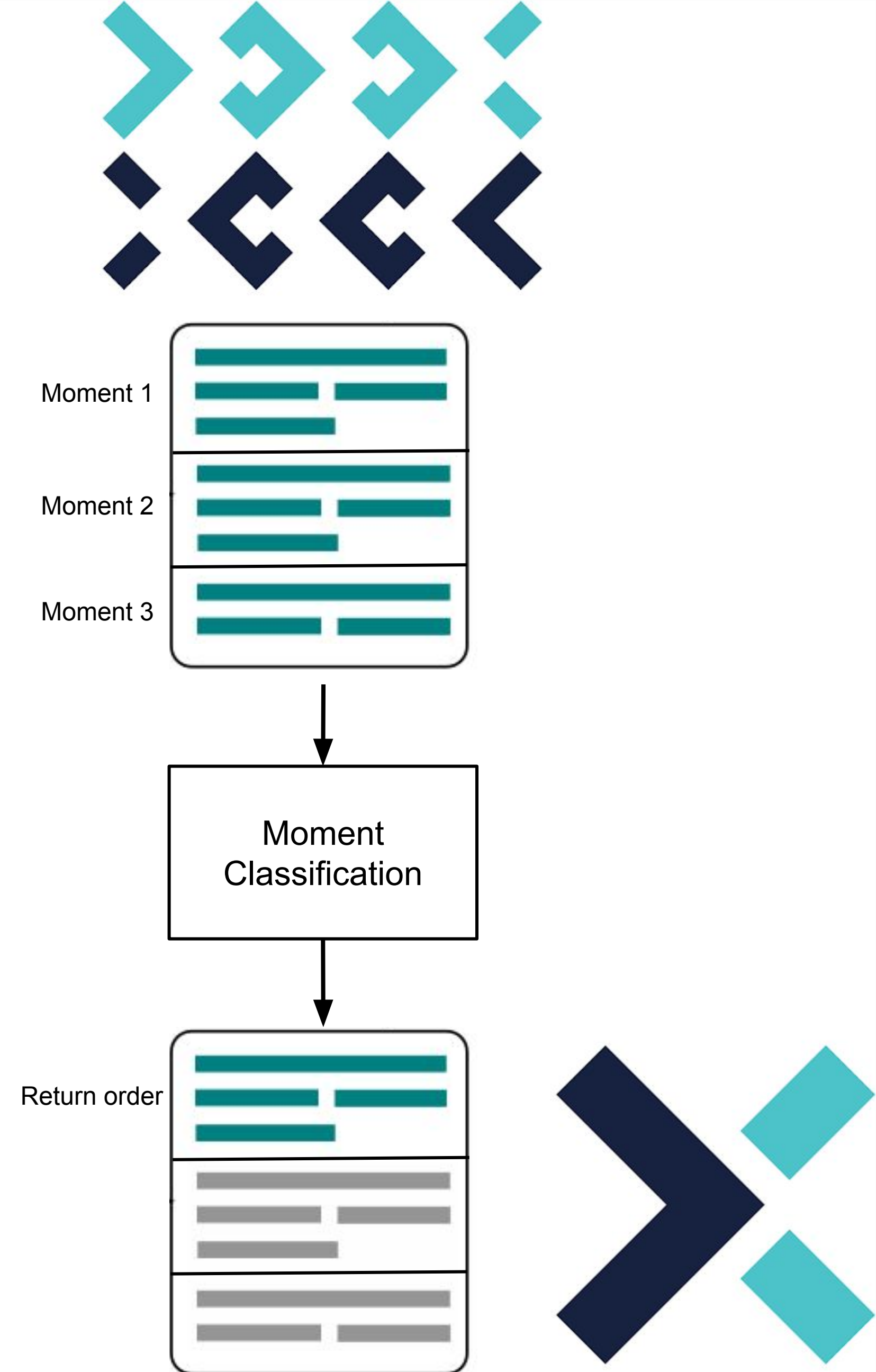


Moment Classification

Contains an algorithm to calculate **embeddings** of moments.

Stores pre-calculated embeddings for **predefined classes** of moments.

For each **moment**, calculate the **embedding** and **classify** it by comparing with the pre-calculated embeddings for different classes.





Approach



Dataset



MultiWOZ

Fully-labeled collection of written conversation spanning over multiple domains and topics.

<https://github.com/budzianowski/multiwoz>



Models

Using embeddings

BERT

<https://github.com/UKPLab/sentence-transformers>

InferSent

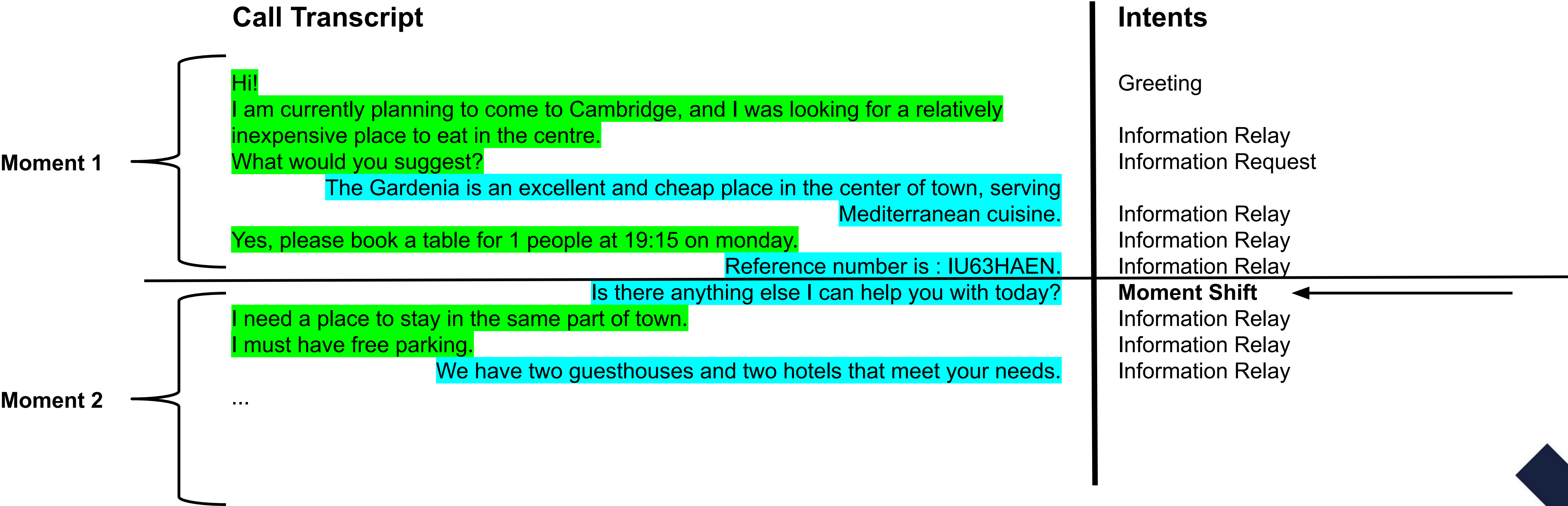
<https://github.com/facebookresearch/InferSent>

Universal Sentence Encoder

<https://tfhub.dev/google/universal-sentence-encoder-large/3>



Real Scenario - Moment Change Detection



Real Scenario - Moment Classification

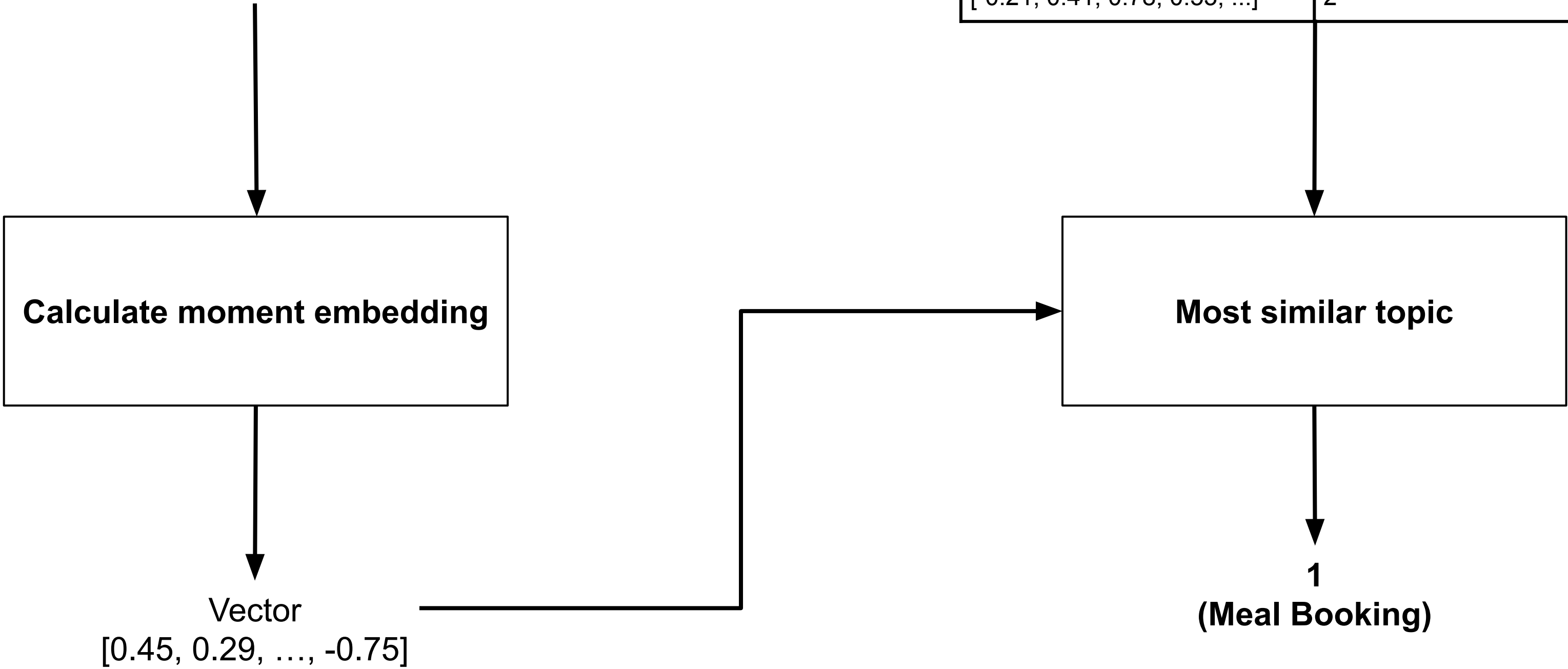


Moment transcript

Hi!
I am currently planning to come to Cambridge, and I was looking for a relatively inexpensive place to eat in the centre.
What would you suggest?
The Gardenia is an excellent and cheap place in the center of town, serving Mediterranean cuisine.
Yes, please book a table for 1 people at 19:15 on monday.
Reference number is : IU63HAEN.

Reference moment embeddings

Moment Embedding	Topic
[0.24, 0.86, 0.12, -0.84, ...]	1
[0.21, -0.32, -0.53, 0.43, ...]	2
[-0.42, 0.32, 0.32, 0.54, ...]	1
[0.54, -0.68, 0.96, -0.12, ...]	4
[-0.67, 0.94, -0.76, -0.42, ...]	3
[-0.21, 0.41, 0.78, 0.53, ...]	2



Results

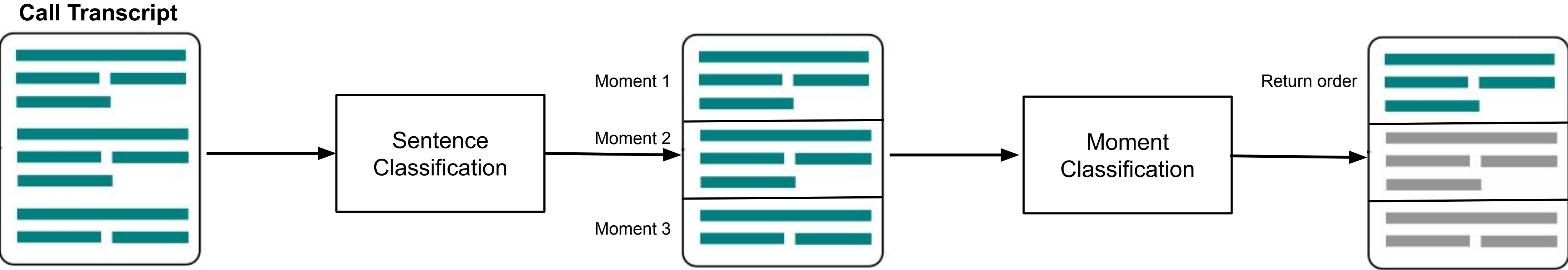


	Accuracy	Prediction time (s)
BERT	83.76%	1.5076
InferSent	70.06%	0.869
Universal Sentence Encoder	82.48%	4.6024
TF-IDF	79.62%	0.22

Using moment embeddings.



Recap



Special Thanks

Diogo Soares Ferreira

diogo.soares@talkdesk.com

Fernando Pais

fernando.pais@talkdesk.com





Thank You!

Pedro Verruma
Engineering Manager



pedroverruma



@pmav



You don't understand embeddings! [0.23, 0.74, 0.12]





¿?

