

Chatbot - Projeto PLN 2020.2

Antônio Paulino
Marcos Gabriel
Pedro Silva

Sumário

- Introdução
- Dataset
- Classificador de intenções
- Extração de entidades
- Demo chatbot

Introdução ao chatbot

De maneira geral, a idéia de se utilizar um chatbot é que ele automatize tarefas específicas de um determinado domínio, e que a interação seja de maneira mais natural possível de forma que mesmo com alguns ruídos de comunicação, como escrita com erro, ainda assim seja possível realizar a tarefa.


Então é importante que o bot reconheça e entenda plenamente a idéia contida em cada frase enviada pelo usuário para que assim possa tomar a decisão mais acertada e resolver o problema do usuário.

Dataset




- Dados de diálogos de clientes com empresas de seguro
- Base **rotulada** com as intenções e as entidades

	conversationId	turnNumber	sentenceNumber	utteranceId	utterance	slot-labels	intent
0	b668a007-9e23-4e4f-8943-e3000cf27efe	1	0	<CONV>b668a007-9e23-4e4f-8943-e3000cf27efe<TUR...	good evening sir	O O O	openinggreeting
1	b668a007-9e23-4e4f-8943-e3000cf27efe	1	1	<CONV>b668a007-9e23-4e4f-8943-e3000cf27efe<TUR...	i need your help	O O O O	outofdomain
2	b668a007-9e23-4e4f-8943-e3000cf27efe	3	0	<CONV>b668a007-9e23-4e4f-8943-e3000cf27efe<TUR...	4523	ssn	contentonly
3	b668a007-9e23-4e4f-8943-e3000cf27efe	4	0	<CONV>b668a007-9e23-4e4f-8943-e3000cf27efe<TUR...	tamilarasi k	name name	contentonly
4	b668a007-9e23-4e4f-8943-e3000cf27efe	6	0	<CONV>b668a007-9e23-4e4f-8943-e3000cf27efe<TUR...	honda city	car_brand O	contentonly




Dataset



[Pull requests](#)
[Issues](#)
[Marketplace](#)
[Explore](#)


[awslabs / multi-domain-goal-oriented-dialogues-dataset](#)

 6
  9
  1

[Code](#)
[Issues 1](#)
[Pull requests](#)
[Actions](#)
[Projects](#)
[Wiki](#)
[Security](#)
[Insights](#)

[master](#)
[1 branch](#)
[0 tags](#)

[Go to file](#)
[Add file](#)
[Code](#)

 **jasonkrone** Update license to CDLA permissive baa3063 on 3 Dec 2020 5 commits

data	Added unlabeled conversations and IC/SL splits from paper	13 months ago
CODE_OF_CONDUCT.md	Initial commit	15 months ago
CONTRIBUTING.md	Initial commit	15 months ago
LICENSE.txt	Update license to CDLA permissive	9 months ago
NOTICE	Initial commit	15 months ago
README.md	Update license to CDLA permissive	9 months ago

About

Data from the publication "Multi-Domain Goal-Oriented Dialogues (MultiDoGO): Strategies toward Curating and Annotating Large Scale Dialogue Data"

[Readme](#)

[View license](#)

Releases

No releases published

Intenções

1. checkclaimstatus
 - a. “please update the my claim status”
2. closinggreeting
3. confirmation
4. contentonly
5. getproofofinsurance
 - a. “i need insurance proof of car”
6. openinggreeting
7. outofdomain
8. rejection
9. reportbrokenphone
 - a. “i want to fix my phone screen”
10. thankyou

Entidades

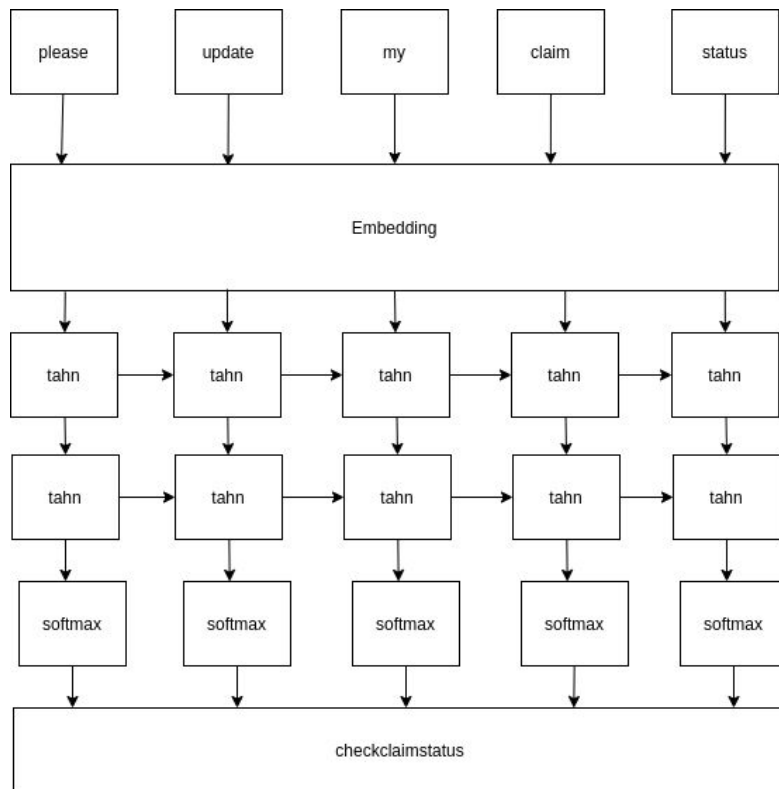
- O
- claimid
- ssnn
- email_address
- policyid
- car_model
- car_brand
- car_year
- name
- phone_number

Classificador de intenção (Intent)

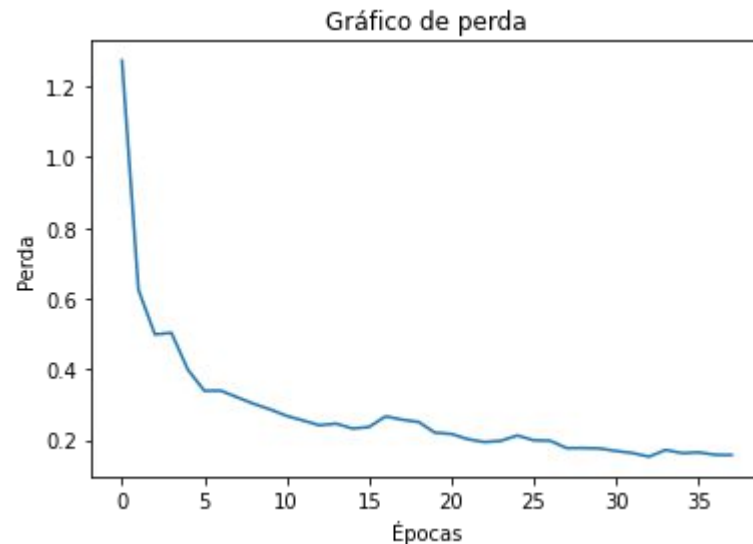
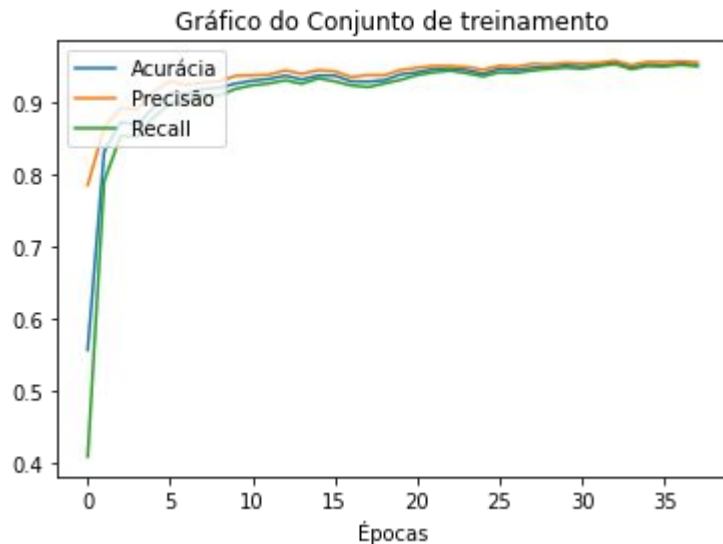
A primeira tarefa a ser realizada pelo bot na sua missão de ajudar o usuário, é entender qual o propósito que o usuário deseja atingir, e categorizar como uma intent. Para determinar qual a intenção do usuário, foi implementada uma rede neural onde dadas as palavras da sentença como por exemplo “What is the status of my claim?” ele deve reconhecer como `search_claim_status` por exemplo. Então, a rede deve ser capaz de abstrair os detalhes da frase e generalizar para identificar corretamente qual o objetivo do usuário.

O classificador de intenções foi construído a partir de uma rede LSTM com duas camadas.

Classificador de intenção (Intent) - Cont. (rede e saída)



Classificador de intenção (Intent) - Cont. (rede e saída)



Classificador de intenção (Intent) - Cont. (rede e saída)

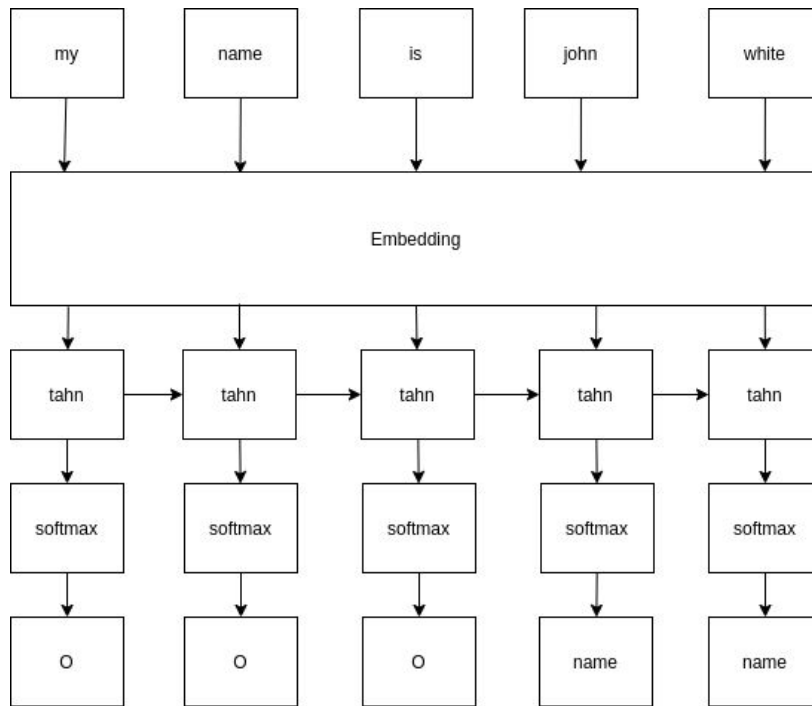
	precision	recall	f1-score	support
0	0.97	0.95	0.96	222
1	0.94	0.97	0.95	260
2	0.90	0.92	0.91	260
3	0.98	0.73	0.84	264
4	1.00	0.95	0.97	264
5	0.97	0.99	0.98	241
6	0.70	0.82	0.76	251
7	0.86	0.95	0.90	278
8	0.95	0.95	0.95	272
9	0.93	0.90	0.92	292
accuracy			0.91	2604
macro avg	0.92	0.91	0.91	2604
weighted avg	0.92	0.91	0.91	2604

Classificador de objeto (entity)

Da mesma forma que o classificador de intenção, o classificador de objeto precisa identificar dentro da frase do usuário qual é o objeto referido (tipo) e quais as suas características se houverem.

O classificador de objetos, assim como o de entidades foi construído a partir de uma rede LSTM, nesse caso com uma camada. Ele categoriza cada palavra da frase com uma tag.

Classificador de objeto (entity)



[illegible]

Dificuldades

- Rasa
- Gerenciamento do Tempo
- Redes neurais

Demo Chatbot