Sebastián Arbeláez Castaño

arbelaezsebastian16@gmail.com

Sales service Chatbox

This proyect was based on a sales site service chatbox[1] that helps customers with various requirements of information. The model takes text input and interpret it to give a response in the different categories of information. The idea was to implement the chatbox but instead of using a feed-forward neural network to interpret the text input a Bert pretrained model was tuned for the task (in this case distilbert was used for the smaller size). The chatbox was correctly implementen in python and was uploaded to a Flask container to be used in javacript.

Index

[1] https://www.youtube.com/watch?v=Da-iHgrmHYg&ab_channel=PythonEngineer

Database structure

To train the model a series of commands was used that associate text with different labels, here are some examples:

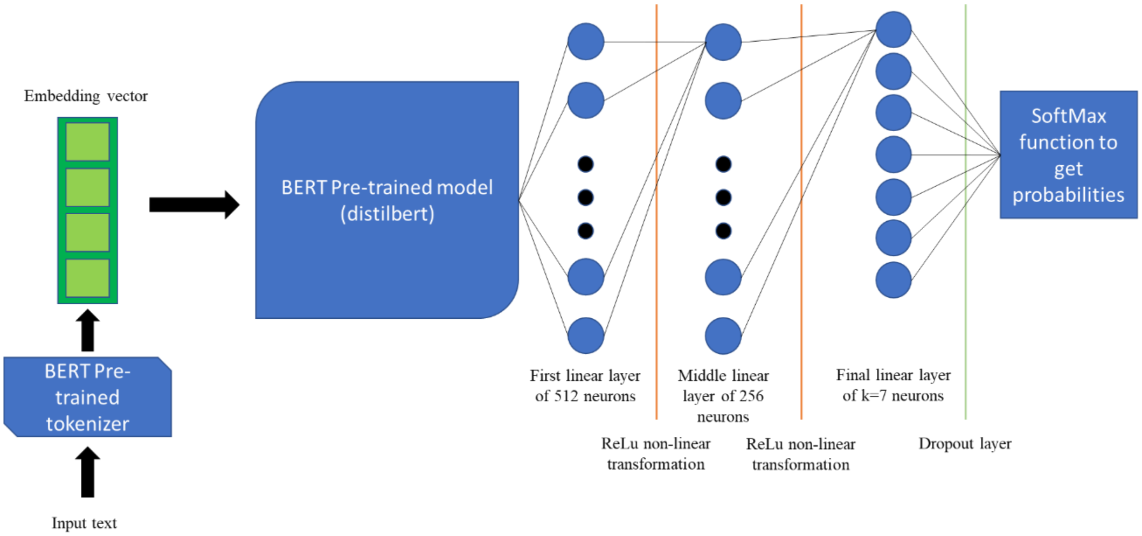
Label	Text
greeting	Hi
greeting	How are you
greeting	Is anyone there?
goodbye	See you later
goodbye	Goodbye
thanks	Thanks
thanks	Thank you
items	What do you sell?
payments	Do you take credit cards?
delivery	When do I get my delivery?

The database was created from a json file containing the different labels, the text associated with them and the possible responses to each label. Here are some examples of the responses:

Label	Response
greeting	Hello, thanks for visiting
greeting	Hi there, what can I do for you?
greeting	Hi there, how can I help?
goodbye	Have a nice day
goodbye	Bye! Come back again soon.
thanks	Any time!
thanks	My pleasure
items	We have coffee and tea
payments	We accept most major credit cards, and Paypal
delivery	Shipping takes 2-4 days

Model Architecture

The model consists on the Bert tokenizer that converts text into Bert vector embedding. This embedding is the processed by the Bert pre-trained model that then gives the output to a feed-forward neural network of three layers, the first receives input of size 768 from Bert, and outputs a vector of 512 values, the next outputs a vector of 256 values and the last layer outputs a vector of the number of classes, in this example there were only 7 classes to decide. Between each linear layer there is a ReLu nonlinear transformation and after the 3 linear layers there is a dropout layer to help in the training step and finally a SoftMax layer to turn the values into probabilities. The category with the highest probability is the chosen by the bot to send a message.


Embedding vector
BERT Pre-trained tokenizer
Input text
BERT Pre-trained model (distilbert)
First linear layer of 512 neurons
Middle linear layer of 256 neurons
Final linear layer of k=7 neurons
ReLu non-linear transformation
ReLu non-linear transformation
Dropout layer
SoftMax function to get probabilities

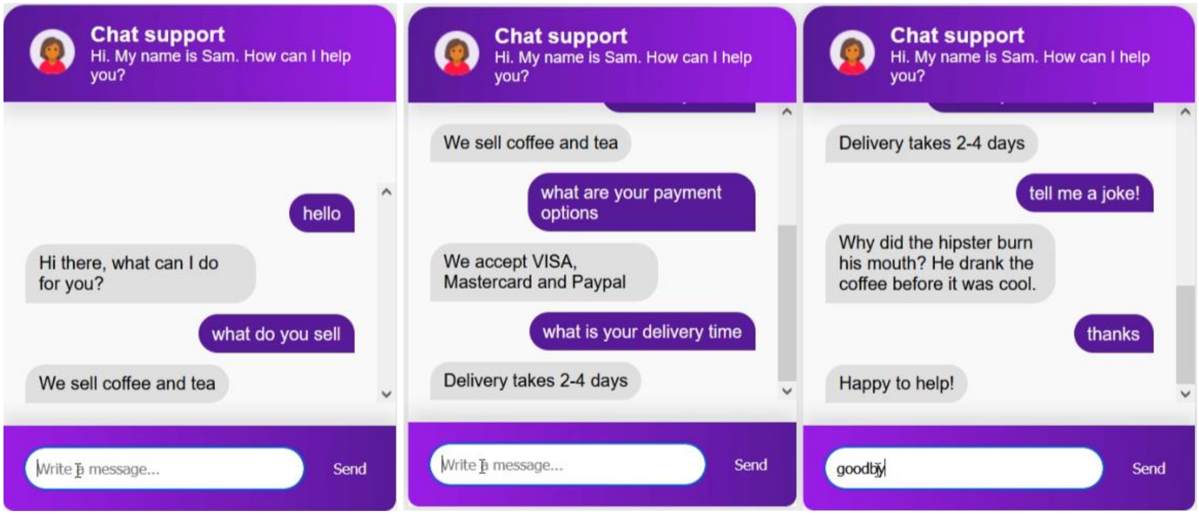
Model tuning

To calibrate the last linear layers after the pretrained BERT the whole dataset was passed thought the network by 200 epochs. The optimizer used was ADAMW from the Transformers library, an optimizer specifically to tune these models. The loss function chosen for this assignment is the negative loss likelihood which tries to maximize the probabilities of the correct category for each entry while minimizing the probability of the others.

$$l(\theta) = -\sum_{i=1}^{n} y_i * \log(\hat{y}_{\theta,i}) + (1 - y_i)\log(1 - \hat{y}_{\theta,i})$$

Results

The model could answer correctly to different texts but in this assignment a small number of examples for each categorie were given. In an extensive example the model should have as more examples as possible to accurately select the category that the persona is giving through the input. Here are some examples of the responses[2].


Chat support
Hi. My name is Sam. How can I help you?
Write a message...
Send

hello

Hi there, what can I do for you?

what do you sell

We sell coffee and tea

Chat support
Hi. My name is Sam. How can I help you?
Write a message...
Send

We sell coffee and tea

what are your payment options

We accept VISA, Mastercard and Paypal

what is your delivery time

Delivery takes 2-4 days

Chat support
Hi. My name is Sam. How can I help you?
Send

Delivery takes 2-4 days

tell me a joke!

Why did the hipster burn his mouth? He drank the coffee before it was cool.

thanks

Happy to help!

goodby

Conclusion

The chatbox is an incredible tool to handle useful information to various costumers without recurring to human capital. The capacity of understanding human language with BERT models is a useful tool to make chatbox more reliable and robust against costumer demands. In the future most of the customer service will be handled by these kind of tools and is important to make them as efficient as possible giving that it will be one of the faces of the company will show to customers.

Bibliography

- BERT chat bot adaptation: https://medium.com/geekculture/simple-chatbot-using-bert-and-pytorch-part-1-2735643e0baa

[2] Please feel free to watch the full video through the hyperlink.

- Chat box implementation with feed-forward neural network with data pipelines: https://www.youtube.com/watch?v=Da-iHgrmHYg&ab_channel=PythonEngineer
- Chat bot integration with javascript: https://www.youtube.com/watch?v=a37BL0stIuM&ab_channel=PythonEngineer