Models comparison

Enrico Ladisa

July 19, 2020

1 Baseline Model

Model used as baseline. The model takes in input the embeddings of the words of the sentence and computes the average.
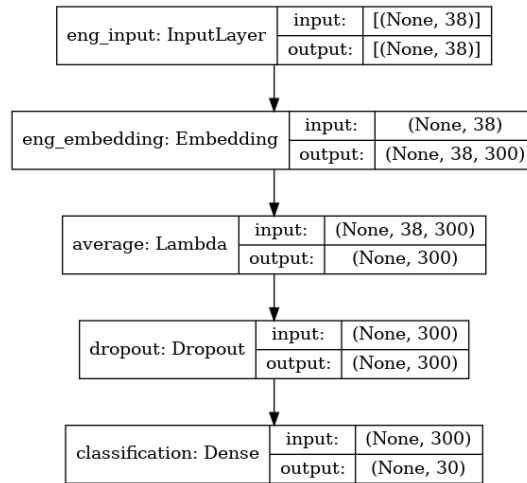

Figure 1: Baseline model architecture.

Set	Loss	Accuracy	Recall	Precision
Training	1.7755	0.5749	0.1334	0.8515
Validation	1.8116	0.5555	0.1272	0.8507
Testing	1.8158	0.5642	0.125	0.8609

2 English Recurrent Model

This model applies a recurrent layer of GRU cells to the list of the embeddings of the sentence.
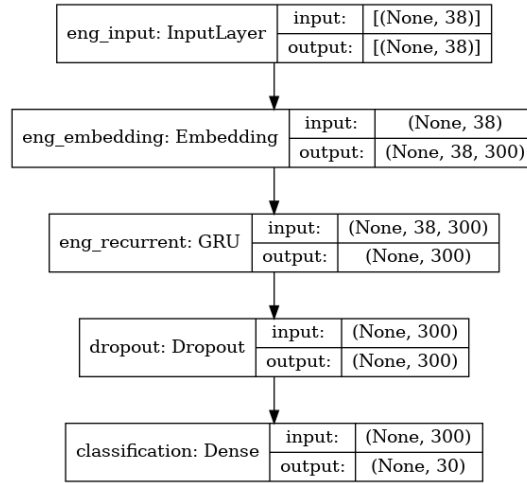
eng_input: InputLayer	input:	[(None, 38)]
	output:	[(None, 38)]

eng_embedding: Embedding	input:	(None, 38)
	output:	(None, 38, 300)

eng_recurrent: GRU	input:	(None, 38, 300)
	output:	(None, 300)

dropout: Dropout	input:	(None, 300)
	output:	(None, 300)

classification: Dense	input:	(None, 300)
	output:	(None, 30)

Figure 2: English model architecture.

Set	Loss	Accuracy	Recall	Precision
Training	0.2805	0.9157	0.8851	0.943
Validation	0.7839	0.7809	0.7372	0.8336
Testing	0.7906	0.7769	0.7356	0.8327

3 Japanese Recurrent Model

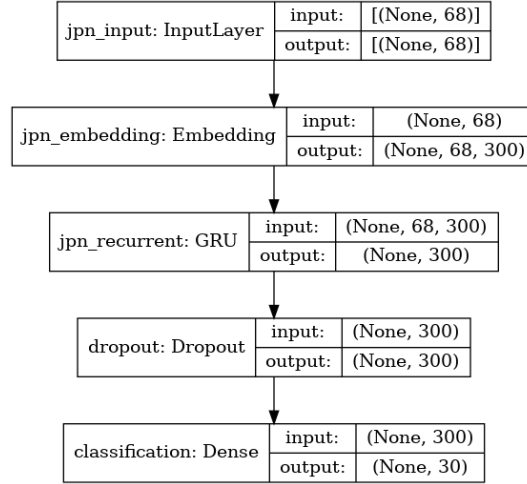This time we apply the same recurrent architecture to the japanese translation of the sentences.

Figure 3: Japanese model architecture.

Set	Loss	Accuracy	Recall	Precision
Training	0.4169	0.8748	0.8253	0.9207
Validation	0.8559	0.7524	0.7003	0.8237
Testing	0.8712	0.752	0.6994	0.823

4   Combined Model

This model combines the outputs of the reccurrent layers of the english and japanese model. The sentence embeddings are first concatenate, then the dimension of the result is reduced to 300 by a Dense layer. This is done so that the sentence embedding to classify has the same dimension of the ones produced by the single-language models.
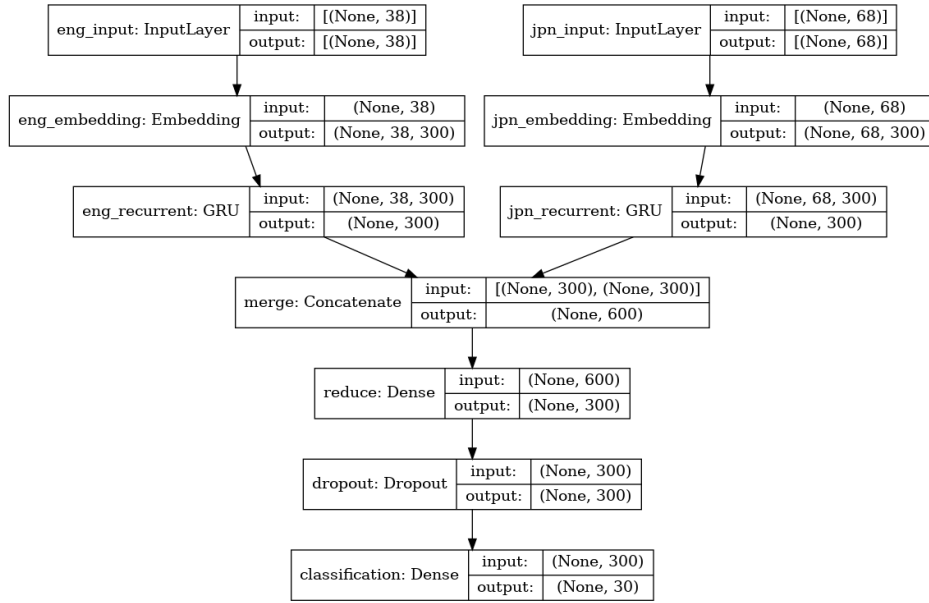

Figure 4: Combined model architecture.

Set	Loss	Accuracy	Recall	Precision
Training	0.1673	0.9485	0.9368	0.9584
Validation	0.5905	0.8367	0.8113	0.8658
Testing	0.5826	0.8363	0.8125	0.8676