

# Multilingual Transformer Ensembles for Portuguese Natural Language Tasks

**Ruan Chaves Rodrigues** ( UFG ) \* - ruanchaves93@gmail.com  
**Jéssica Rodrigues da Silva** ( UFSCar ) \*\* - jsc.rodrigues@gmail.com  
**Pedro Vitor Quinta de Castro** ( UFG ) \* - pedrovitorquinta@inf.ufg.br  
**Nádia Félix Felipe da Silva** ( UFG ) \* - nadia@inf.ufg.br  
**Anderson da Silva Soares** ( UFG ) \* - anderson@inf.ufg.br

\* : Institute of Informatics

Federal University of Goiás ( UFG ), Brazil

\*\* : Department of Computer Science

Federal University of São Carlos ( UFSCar ), Brazil

March 2, 2020



**DEEP LEARNING**  
**BRASIL**

# Agenda

- 1 Introduction
- 2 Related Work
- 3 Architecture
- 4 Results
- 5 Conclusions
- 6 References

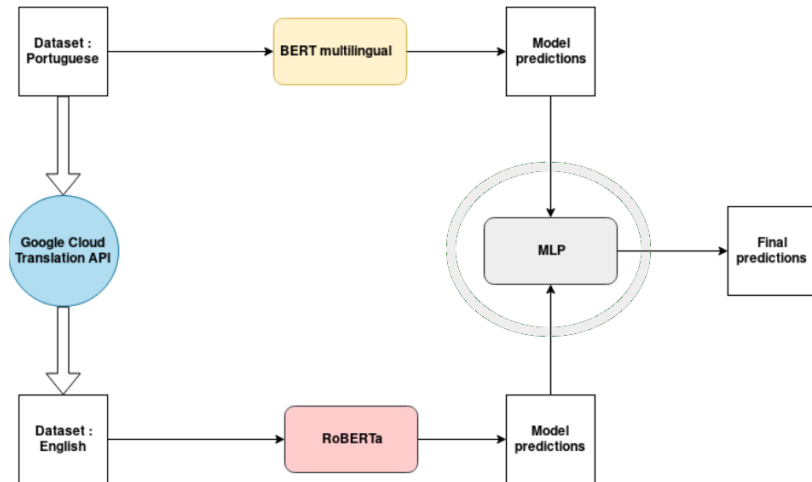
- Semantic Similarity Tasks on Low-Resource Languages

Strategy	Trade-off
Language models pretrained on the low-resource language	Expensive to train and easy to deploy, but behind the state-of-the-art.
Train your own task-specific translation model and leverage annotated data from high-resource languages ( Tang et al. <a href="#">[2018]</a> )	Good results, but potentially expensive.
<b>Translate the dataset through an external API and leverage models pretrained on high-resource languages</b>	<b>Best possible results, but limited scalability.</b>

# Related Work

- BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding ( Devlin et al. [\[2018\]](#) )
- RoBERTa: A Robustly Optimized BERT Pretraining Approach ( Liu et al. [\[2019\]](#) )

# Architecture



# A Hack

[ 0.345 -0.134 1.128 0.845 ... 0.012 ]



[ 0 0 1 1 ... 0 ]



[ None, None, Entailment, Entailment, ..., None ]

# Results

ASSIN 2	Portuguese	Semantic Textual Similarity		Recognizing Textual Entailment	
	Sources	PCC	MSE	Acc	F1
Teams	Deep Learning Brasil	0.79	0.59	<b>0.88</b>	<b>0.88</b>
	Stilingue	0.80	<b>0.39</b>	0.87	0.87
	IPR	0.83	<b>0.52</b>	0.88	0.88
Models	BERT-multilingual	0.77	1.22	0.81	0.80
	portuguese-BERT	<b>0.84</b>	0.73	0.83	0.82
	RoBERTa	<b>0.85</b>	0.74	<b>0.89</b>	<b>0.89</b>
	Word Overlap Baseline	0.58	0.75	0.67	0.68

# Results

ASSIN 2							
				Task			
				Semantic Textual Similarity		Recognizing Textual Entailment	
English model	Portuguese model	Ensemble	Folds	PCC	MSE	Acc	F1
RoBERTa-large	BERT-multilingual	averaging		<b>0.86</b>	0.91	0.88	0.88
		stacking	5	0.79	0.59	0.88	0.88
			10	<b>0.85</b>	<b>0.50</b>	<b>0.89</b>	<b>0.89</b>
			20	0.85	0.50	0.89	0.89
	portuguese-BERT	averaging		0.86	0.71	0.89	0.89
		stacking	5	0.80	0.68	0.89	0.89
			10	0.86	0.51	0.89	0.89
			20	<b>0.86</b>	<b>0.45</b>	<b>0.89</b>	<b>0.89</b>
			32	0.86	0.49	0.89	0.89



# Results

ASSIN 1 ( Brazilian Portuguese )							
				Task			
		Ensemble	Folds	Semantic Textual Similarity		Recognizing Textual Entailment	
English model	Portuguese model			PCC	MSE	Acc	F1
RoBERTa-large	BERT-multilingual	averaging		0.86	0.22	0.83	0.46
		stacking	5	0.86	0.22	0.78	0.30
			10	<b>0.86</b>	<b>0.21</b>	<b>0.89</b>	0.55
			20	0.86	0.22	0.89	<b>0.56</b>
		averaging		0.87	0.20	0.80	0.36
	portuguese-BERT	stacking	5	0.87	0.24	0.78	0.32
			10	0.87	0.19	0.88	<b>0.56</b>
			20	<b>0.87</b>	<b>0.19</b>	<b>0.89</b>	0.55
		averaging					

# Results

ASSIN 1 ( European Portuguese )							
				Task			
				Semantic Textual Similarity		Recognizing Textual Entailment	
English model	Portuguese model	Ensemble	Folds	PCC	MSE	Acc	F1
RoBERTa-large	BERT-multilingual	averaging		0.88	0.42	0.80	0.49
		stacking	5	0.88	0.40	0.88	0.58
			10	<b>0.88</b>	<b>0.36</b>	<b>0.88</b>	<b>0.58</b>
			20	0.88	0.38	0.88	0.58
		averaging		0.88	0.39	0.81	0.50
	portuguese-BERT	stacking	5	0.88	0.33	0.88	0.58
			10	<b>0.88</b>	<b>0.33</b>	<b>0.88</b>	<b>0.58</b>
			20	0.88	0.37	0.88	0.58
		averaging		0.88	0.39	0.81	0.50

# Analysis

t	h	score	gold_score	error	
0	Uma pessoa está andando de bicicleta ao longo de uma trilha enlameada	Um homem está andando de <b>mountain bike</b> descendo um morro de terra	4.30	1.75	2.55
1	Um skatista está dando um pulo de uma rampa	Uma mulher vestida de preto está usando patins e executando um truque em um trilho	3.50	1.00	2.50
2	A menina de uniforme azul e branco está torcendo	Alguns números coreográficos estão sendo apresentados pela menina de uniforme azul e branco	3.86	1.50	2.36
3	Um cachorro vermelho está pulando e pegando uma bola de tênis <b>na boca</b>	Um cachorro caramelo está pulando para cima e pegando uma bola de tênis	4.31	2.00	2.31
4	Uma menina pequenina <b>de brinquedo</b> está em um carrinho de brinquedo	Uma criança está sentada em um carro de brinquedo vermelho e amarelo	4.03	1.75	2.28
5	<b>Um cachorro preto e um branco</b> estão correndo alegremente na grama	<b>Uma pessoa negra vestindo branco</b> está correndo alegremente com o cachorro na grama	3.69	1.50	2.19
6	Um homem velho e uma mulher estão andando de bicicleta no calçadão perto do oceano	Um homem e uma mulher de meia idade estão montando bicicletas por uma estrada ao lado da água	4.19	2.00	2.19
7	Esse grupo de pessoas está praticando salvamento na água e usando coletes salva-vidas	A corrente está sendo navegada por um grupo de amigos em um bote	3.43	1.25	2.18
8	Uma mulher está praticando <b>snowboard</b> ladeira abaixo e outra mulher está de pé atrás dela	Um menino loiro está deslizando por uma duna de areia em um <b>skate</b>	3.27	1.10	2.17
10	Dois meninos estão fazendo artes marciais em um tapete azul	Duas crianças de roupas brancas e equipamento de segurança vermelho estão brigando em um tapete	3.87	1.75	2.12
11	Papel e tesoura ambos cortam	O papel está sendo cortado com tesoura	3.61	1.50	2.11

# Analysis

	t	h	score	gold_score	error
2456	Um cachorro está reunindo um bando de ovelhas <b>brancas em uma fazenda empoeirada</b>	Um cão pastor está dispersando um rebanho de ovelhas	3.76	4.40	-0.64
2457	Um pato grande está voando acima de um riacho rochoso	<b>Não</b> tem <b>nenhuma</b> ave marinha em cima de algumas rochas	3.16	3.80	-0.64
2458	Um homem está quebrando uma laje de concreto com uma marreta	<b>Não</b> tem <b>nenhum</b> homem quebrando uma laje de concreto com uma marreta	3.83	4.50	-0.67
2459	Um homem negro está <b>dançando break</b> na calçada	O homem de pele escura está sobre uma mão em frente a um prédio amarelo	3.22	3.90	-0.68
2460	Um homem em uma motocicleta <b>não</b> está andando em uma roda	Um homem em uma motocicleta está andando sobre uma roda	3.78	4.50	-0.72
2461	Um chefe está <b>preparando alguma comida</b>	Um chefe está <b>jogando fora algum alimento</b>	3.07	3.80	-0.73
2462	O palhaço está falando <b>no microfone</b>	Uma pessoa está falando	3.00	3.75	-0.75
2463	Um cavalo e seu cavaleiro <b>não</b> estão saltando sobre um obstáculo	Um cavalo está pulando o obstáculo e tem um cavaleiro <b>nas suas costas</b>	3.81	4.60	-0.79
2464	O homem está fazendo exercícios	Dois homens estão <b>lutando boxe</b>	1.99	2.80	-0.81
2465	Uma cebola está sendo fatiada pela senhora	A senhora está fatiando uma cebola	4.16	5.00	-0.84
2466	Batatas estão sendo fatiadas por um cara	O cara está fatiando a batata	3.36	4.25	-0.89
2467	Está chovendo em um homem a pé	Um homem está andando na chuva	3.81	4.90	-1.09
2468	Uma batata está sendo fatiada pelo cara	O cara está fatiando a batata	3.63	4.75	-1.12
2469	<b>O</b> cara está fatiando <b>a</b> batata	<b>Um</b> cara está fatiando <b>uma</b> batata	3.52	5.00	-1.48
2470	O cara está fatiando <b>uma</b> batata	O cara está fatiando <b>a</b> batata	3.51	5.00	-1.49
2471	Uma pessoa está se exercitando	Um homem está dançando	1.85	3.38	-1.54

# Conclusions

- Our model has potentially severe limitations for dealing with negatives, double negatives, articles, loanwords, grammatical voice and distinctions that rely primarily on semantic roles. (Kassner and Schütze [2019] , Ettinger [2019])
- Solutions:
  - Add word embeddings and traditional NLP hand-crafted features (Tian et al. [2017]).
  - Add BERT-generated features:
    - **BERT for Negation Detection** : NegBERT: A Transfer Learning Approach for Negation Detection and Scope Resolution Khandelwal and Sawant [2019]
    - **BERT for Paraphrase Detection** : "PAWS: Paraphrase Adversaries from Word Scrambling" Zhang et al. [2019]
- **Source code:** <https://github.com/ruanchaves/assin>

# References I

- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding.
- Ettinger, A. (2019). What bert is not: Lessons from a new suite of psycholinguistic diagnostics for language models.
- Kassner, N. and Schütze, H. (2019). Negated lama: Birds cannot fly.
- Khandelwal, A. and Sawant, S. (2019). Negbert: A transfer learning approach for negation detection and scope resolution.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach.
- Tang, X., Cheng, S., Do, L., Min, Z., Ji, F., Yu, H., Zhang, J., and Chen, H. (2018). Improving multilingual semantic textual similarity with shared sentence encoder for low-resource languages.
- Tian, J., Zhou, Z., Lan, M., and Wu, Y. (2017). ECNU at SemEval-2017 task 1: Leverage kernel-based traditional NLP features and neural networks to build a universal model for multilingual and cross-lingual semantic textual similarity. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, Vancouver, Canada. Association for Computational Linguistics.
- Zhang, Y., Baldridge, J., and He, L. (2019). Paws: Paraphrase adversaries from word scrambling.

# Acknowledgements



**DEEP LEARNING**  
**BRASIL**



**INSTITUTO DE**  
**INFORMÁTICA**  
**UFG**

# Acknowledgements

