



UM MODELO DE TRADUÇÃO AUTOMÁTICA NEURAL PARA PORTUGUÊS-LIBRAS

Marcos A. H. Boriola¹, Gustavo H. Paetzold² e Aline K. Pin²

O acesso à informação de forma independente por pessoas surdas ainda é muito dificultado devido ao uso único da Libras como forma de comunicação, já que a mesma é uma língua visuoespacial. Porém existe uma versão das línguas de sinais em forma escrita conhecida como Sutton Sign-Writing (SW). A SW representa cada sinal em figuras com as configurações necessárias para sua realização. Tomando como iniciativa a necessidade de aumentar a acessibilidade das pessoas surdas às informações, o objetivo deste trabalho é a criação de um tradutor de Português para Libras na forma SW. Para realização deste trabalho foi desenvolvido uma rede neural utilizando o modelo Seq2Seq. A rede neural possui apenas uma camada oculta, que foi testada com 64, 128, 256 e 512 nós. Ademais, a mesma foi treinada com o número de iterações entre 10 a 100, a cada 10, 200 e 1000, a cada 100 e de 2000 a 100000, a cada 1000. A rede recebe como entrada uma frase em Português e tem como saída uma frase em Libras na forma SW. Para a etapa de treinamento e validação utilizouse dados retirados da Literatura Brasileira do site SignPuddle. Como parte da análise dos resultados foram usadas as métricas WER e BLEU. Para a métrica WER, teve-se uma média entre as iterações de 2,35; 1,32; 2,11 e 4,08 para 64, 128, 256 e 512 nós. Um bom tradutor moderno obtém bem menos de 1,0. Já para a BLEU, as médias para todos os nós foi 0, onde 0 é o pior valor e 1 o melhor. Em relação a outros tradutores que utilizam do mesmo método, este modelo apresentou um resultado insatisfatório devido à escassez de dados para o treinamento. Isso porque, em geral, são empregadas milhões de instâncias para o treinamento, enquanto tínhamos apenas centenas.

Palavras-chave: Português; Libras; Escrita de sinais; Rede neural; Tradutor.

¹Acadêmico(a) do Curso de Engenharia da Computação.

² Orientador(a).

UM MODELO DE TRADUÇÃO AUTOMÁTICA NEURAL PARA PORTUGUÊS-LIBRAS

Marcos A. H. Boriola¹; Gustavo H. Paetzold²; Aline K. Pin³

¹Acadêmico do Curso de Engenharia de Computação, Universidade Tecnológica Federal do Paraná – Toledo, mahboriola@gmail.com;

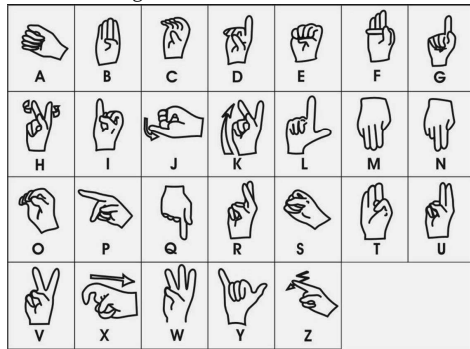
²Orientador, COECO, Nome da Universidade – Universidade Tecnológica Federal do Paraná – Toledo, ghpaetzold@utfpr.edu.br;

³Orientadora, COMAT, Nome da Universidade – Universidade Tecnológica Federal do Paraná – Toledo, alinepin1810@gmail.com.

INTRODUÇÃO

O acesso à informação de forma independente por pessoas surdas ainda é muito dificultado devido ao uso da Libras (Língua Brasileira de Sinais) como forma de comunicação, já que a mesma é uma língua visuoespacial.

Figura 1 – Alfabeto em Libras



Fonte: super.abril.com.br/mundo-estranho/quem-criou-a-linguagem-de-sinais-para-surdos/

Porém existe uma versão das línguas de sinais em forma escrita conhecida como *Sutton SignWriting* (SW). A SW representa cada sinal em figuras com as configurações necessárias para sua realização.

Figura 2 – Alfabeto em SW. Adaptado.



Fonte: www.signwriting.org

E na tentativa de sanar este problema este trabalho propõem um tradutor de Português para Libras na forma SW.

SOLUÇÃO

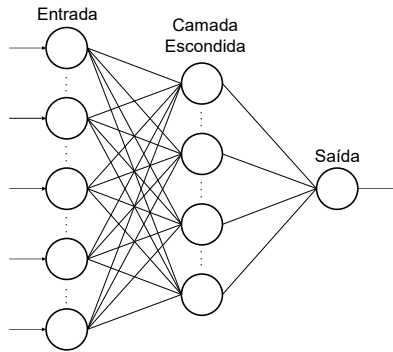
Para realização deste trabalho foi desenvolvida uma rede neural utilizando o modelo *Sequence to Sequence*, mais conhecido como Seq2Seq. Onde a mesma possui uma etapa de codificação e decodificação dos dados de entrada.

Os dados utilizados para as etapas de treinamento e validação foram retirados da Literatura Brasileira da plataforma *SignPuddle* no site *SignBank*.

No total eram 187 instâncias, sendo elas separadas em 150 para o treinamento e os outros 37 para a etapa de validação.

Pela pouca quantidade de dados disponíveis para treinamento, a rede neural desenvolvida possui apenas uma camada oculta, que foi testada com 64, 128, 256 e 512 nós.

Figura 3 – Rede Neural



Fonte: Próprio autor

A rede recebe como entrada uma frase em Português e tem como saída uma frase em Libras na forma SW em sua configuração FSW (*Formal SignWriting*), que é uma representação em formato de código da escrita de sinais.

Figura 4 – Frase “eu gosto de comer doces.” em SW



Fonte: www.signbank.org/signpuddle2.0

A frase demonstrada na figura 4 na configuração FSW se dá pelo código:

M515x511S10b11485x490
M519x516S15a02481x485S21100506x502
M522x556S33b00482x483S15a00495x529S2
2114496x520
M534x560S33b00482x483S15a02493x523S2
e502507x537 S38800464x496.

Ademais, a mesma foi treinada com o número de iterações entre 10 a 100, a cada 10, 200 e 1000, a cada 100 e de 2000 a 100000, a cada 1000.

RESULTADOS

Como forma de análise dos resultados foram utilizadas métricas para comparar a tradução obtida na saída da rede neural com a frase correta em SW, e as métricas foram WER (*Word Error Rate* – Taxa de Erro de Palavras), descrito na equação (1) – sendo S o número de substituições, D o número de deleções, I o número de inserções e C o número de palavras corretas – e BLEU (*Bilingual Evaluation Understudy* – Estudo de Avaliação Bilingue), descrito na equação (2) – sendo P a precisão, M o número de palavras da tradução que são encontradas na frase correta e W_i o total de palavras da tradução.

$$WER = \frac{S + D + I}{S + D + C} \quad (1)$$

$$P = \frac{M}{W_i} \quad (2)$$

Para a métrica WER, teve-se uma média entre as iterações de 2,35; 1,32; 2,11 e 4,08 para 64, 128, 256 e 512 nós. Um bom tradutor moderno obtém bem menos de 1,0. Já para a BLEU, as médias para todos os nós foi 0, onde 0 é o pior valor e 1 o melhor.

Tabela 1 – Média de resultados

Métricas	64	128	256	512
WER	2,35	1,32	2,11	4,08
BLEU	0	0	0	0

Fonte: Próprio autor

CONCLUSÃO

Em relação a outros tradutores que utilizam do mesmo método, este modelo apresentou um resultado insatisfatório devido à escassez de dados para o treinamento. Isso porque, em geral, são empregadas milhões de instâncias para o treinamento, enquanto tínhamos apenas centenas.

Essa baixa quantidade de dados fez com que nosso modelo caísse facilmente num problema de aprendizado de máquina chamado *underfitting*, que ocorre quando a rede neural não consegue reconhecer muito bem os padrões dos dados apresentados, sendo uma dessas causas a baixa quantidade de dados. Portanto, o diminuto número de dados além de trazer problemas com perda de aprendizado do modelo ainda traz o problema de vocabulário reduzido, sendo assim este modelo de tradução está restrito a uma pequena quantidade de palavras presentes nos dados de treinamento.