Nome: Diogo Castanho Emídio
Número USP: 11297274

TraduSino: Tradutor Automático e Inteligente de Português-Chinês

1. Resumo

Este projeto tem como objetivo desenvolver um tradutor automático avançado que facilite a comunicação entre falantes de português e chinês: o TraduSino. Através do uso de técnicas de Processamento de Linguagem Natural (PLN) e Aprendizado de Máquina (AM), visa-se melhorar a eficiência e a qualidade das traduções de textos e, futuramente, áudios. Por fim, também anseia parcerias com instituições de ensino e pesquisa, profissionais qualificados, e empresas de tecnologia, para se obter recursos adicionais e validar o desempenho do sistema.

Palavras-chave: Tradutor Automático. Chinês. Português. Processamento de Linguagem Natural.

2. Contextualização

A China está se tornando uma potência em Inteligência Artificial (IA), e é importante entender o porquê. Desde a década de 50, a IA tem evoluído rapidamente, e já está impactando diversas áreas do nosso dia a dia, incluindo a produção de alimentos, a gestão de energia elétrica, e até mesmo o campo militar. Dito isso, o país entrou no novo milênio como um dos principais atores da economia global e no desenvolvimento de novas tecnologias, saindo de um país agrícola para um modelo em produção industrial e tecnológica.

Em 2015, a China anunciou o plano "Made in China 2025", com o objetivo de se tornar líder global em IA. O programa, que ainda está em vigor, possui dez setores principais a serem investidos, incluindo dois focados apenas em IA e robótica. O governo chinês se tornou um dos que mais investem em pesquisa, desenvolvimento e aplicação de IA no cotidiano de sua população. Empresas como Alibaba, Baidu, SenseTime, Tencente e iFlytek são as maiores empresas no ramo atuando no país.

A nação chinesa está avançando rapidamente na evolução de IAs, sendo protagonista em quantidade de patentes, produção científica, e número de especialistas e companhias focadas no ramo. No entanto, a tensão entre China e Estados Unidos é evidente, com disputas militares e comerciais, tudo para não ficar atrás nessa corrida tecnológica. Além disso, a China é conhecida por ser um dos países com maior nível de vigilância civil do mundo, também com aplicação de IAs.

3. Motivação e Justificativa

A demanda por tradução automática de qualidade entre o português e o chinês tem aumentado devido ao crescente intercâmbio cultural, econômico e acadêmico entre o Brasil e a China, além da ascendente influência dessa nos últimos anos. No entanto, as soluções existentes ainda enfrentam desafios significativos em relação a precisão e fluência das traduções, principalmente em contextos mais complexos e técnicos. Isso também se dá pelo fato de muitas plataformas utilizarem o inglês como intermediário, o que pode agravar a quantidade de informações perdidas no processo. Portanto, o desenvolvimento do presente projeto, batizado de TraduSino, visa preencher essas lacunas, ao ser um tradutor automático robusto, confiável e direto.

4. Objetivos

4.1. Objetivo geral

O objetivo deste projeto é desenvolver um tradutor automático avançado que facilite a comunicação entre falantes de português e chinês, permitindo uma conversão rápida e precisa de textos e, posteriormente, áudio. Assim, busca-se melhorar a eficiência e a qualidade da tradução entre esses dois idiomas, facilitando a interação e promovendo a troca cultural.

4.2. Objetivos específicos

1. Desenvolver uma plataforma online para tradução automática de textos do português para o chinês e vice-versa;
2. Implementar tecnologias avançadas de PLN e AM para melhorar o desempenho das traduções;
3. Pesquisar e desenvolver algoritmos para tradução automática de áudio em tempo real, possibilitando a comunicação oral entre falantes de português e chinês (mais especificamente, mandarim);
4. Realizar testes e avaliações constantes para aprimorar a qualidade da plataforma;
5. Disponibilizar o TraduSino como uma ferramenta gratuita e acessível para uso público, tanto como página web quanto como aplicativo.

5. Metodologia

1. Coleta e compilação de um *corpus* bilíngue extenso de textos nas línguas portuguesa e chinesa para treinamento e aprimoramento dos algoritmos de tradução;
2. Utilização de técnicas de PLN e AM para desenvolver modelos de tradução automática;
3. Desenvolvimento de uma interface intuitiva e amigável para permitir que os usuários traduzam textos de forma fácil e eficiente;
4. Implementação de algoritmos de reconhecimento e síntese de fala para a tradução automática de áudio;
5. Realização de testes e avaliações sistemáticas, incluindo comparações com tradutores automáticos já existentes, para medir a precisão, a fluência e a eficácia do TraduSino.

6. Possíveis Colaborações

Para o desenvolvimento do projeto, cogita-se cooperações com instituições de ensino e pesquisa que possam fornecer recursos adicionais, como acesso a bancos de dados linguísticos e especialistas em PLN. Além disso, pensa-se em colaborações com tradutores profissionais e falantes nativos de português e chinês para aperfeiçoar o sistema e validar seu desempenho. Por fim, pretende-se propor parcerias a gigantes da tecnologia que já possuem tradutores automáticos, como Google e Baidu.

7. Conclusão

Em decorrência de uma definição clara de objetivos e metodologia, o atual projeto tem o potencial de facilitar a comunicação entre falantes de português e chinês, superando os desafios existentes na tradução automática como existe hoje. E ao oferecer uma plataforma avançada e acessível, com possibilidade de contribuições de diversas magnitudes e envolvendo várias entidades e indivíduos, o TraduSino pode contribuir para o intercâmbio entre o Brasil e a China, promovendo a troca de conhecimentos e enriquecendo as relações entre ambos os países.

8. Bibliografia

- Brownlee, J. (2019). Deep Learning for Natural Language Processing. Machine Learning Mastery.
- Koehn, P. (2010). Statistical Machine Translation. Cambridge University Press.
- Sutskever, I., Vinyals, O., & Le, Q. (2014). Sequence to Sequence Learning with Neural Networks. Advances in Neural Information Processing Systems.
- Vaswani, A., et al. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems.
- Luong, M. T., Pham, H., & Manning, C. D. (2015). Effective Approaches to Attention-based Neural Machine Translation. Proceedings of the Conference on Empirical Methods in Natural Language Processing.
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate. arXiv preprint arXiv:1409.047 3.
- Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics.
- Xu, Y., et al. (2021). Transformer-based Speech Translation: A Survey. arXiv preprint arXiv:2101.00671.