



IA INCLUSIVA: APRIMORAMENTO DE LEGENDA AUTOMÁTICAS COM FOCO EM ACESSIBILIDADE AUDITIVA

Andrey Fogaça¹, Flávia Sayuri Arakawa², Sabine Cassol³

¹Acadêmico do Curso de Análise e Desenvolvimento de Sistemas, Campus Ponta Grossa-PR, Universidade Cesumar - UNICESUMAR. Bolsista PIBIC/ICETI-UniCesumar. andreyfogaça1234@gmail.com

²Orientadora, Doutora, Docente no Curso de Engenharia de Software, Universidade Cesumar, UNICESUMAR, Campus de Ponta Grossa- PR. Pesquisadora do Instituto Cesumar de Ciência, Tecnologia e Inovação – ICETI. flavia.arakawa@unicesumar.edu.br

³Especialista, Docente no Curso de Engenharia de Software e Análise e Desenvolvimento de Sistemas, Universidade Cesumar, Unicesumar, Campus Ponta Grossa - PR. sabine.cassol@unicesumar.edu.br

RESUMO

O avanço das tecnologias digitais e o crescimento do consumo de vídeos online evidenciam a importância da acessibilidade comunicacional, especialmente para pessoas com deficiência auditiva. Apesar da popularização das ferramentas de Inteligência Artificial (IA), as legendas automáticas geradas ainda apresentam falhas significativas, sobretudo ao lidar com a variedade linguística e regional do português brasileiro. Este projeto de pesquisa tem como objetivo analisar e propor melhorias na geração de legendas automáticas por meio do uso de técnicas de Processamento de Linguagem Natural (PLN), com ênfase na compreensão contextual por meio de modelos de redes neurais recorrentes, como as LSTM (Long Short-Term Memory). A pesquisa será desenvolvida com abordagem qualitativa, exploratória e de natureza aplicada. A metodologia inclui uma revisão bibliográfica sistemática sobre IA, PLN, acessibilidade e redes neurais, bem como a coleta e análise de dados empíricos por meio de vídeos com legendas automáticas geradas em plataformas como YouTube e TikTok. Os dados serão avaliados manualmente, classificando os principais erros de transcrição e compreendendo os padrões de falhas. Com base nisso, será proposto o esboço teórico de um modelo de IA mais eficiente, capaz de respeitar variações linguísticas e promover maior inclusão comunicacional. A pesquisa busca contribuir tanto no campo teórico, ao aprofundar a discussão sobre acessibilidade e tecnologias inclusivas, quanto no campo prático, oferecendo subsídios para o aprimoramento de sistemas de legenda automática e ampliando as possibilidades de participação social de pessoas com deficiência auditiva. A proposta também está alinhada à Lei Brasileira de Inclusão (LBI) e reforça o compromisso com a democratização da informação e a promoção da equidade no acesso a conteúdos digitais.

PALAVRAS-CHAVE: Acessibilidade digital; Inteligência Artificial; Processamento de Linguagem Natural.

1 INTRODUÇÃO

Com o crescimento contínuo no consumo de conteúdos digitais, a acessibilidade se torna cada vez mais essencial para garantir uma experiência inclusiva a todos os usuários. No entanto, mesmo com os avanços significativos nas tecnologias baseadas em Inteligência Artificial (IA), ainda são evidentes as limitações das ferramentas existentes no que diz respeito à acessibilidade, especialmente para pessoas com deficiência auditiva. A geração automática de legendas, amplamente utilizada em plataformas de vídeo e ambientes virtuais, muitas vezes falha em oferecer precisão semântica e contextual, prejudicando a compreensão por parte desse público.

Diante desse cenário, surge a necessidade de desenvolver soluções mais eficazes que ampliem a inclusão comunicacional. Um dos caminhos promissores para esse avanço está no uso de técnicas de Processamento de Linguagem Natural (PLN), capazes de aprimorar a capacidade das IAs de compreender e transcrever a fala de maneira mais contextualizada. Entre essas técnicas, destacam-se as Redes Neurais Recorrentes (RNNs), que analisam sequências de dados considerando a dependência temporal entre palavras e frases, o que é essencial para a geração de legendas que reflitam com mais fidelidade a intenção e o significado do discurso.



Especificamente, as variantes das RNNs conhecidas como LSTM (Long Short-Term Memory) têm se mostrado altamente eficazes na tarefa de reconhecimento de fala. Essas redes são capazes de "memorizar" informações ao longo de uma sequência de tempo, permitindo que o sistema identifique nuances de linguagem, entonação, regionalismos e padrões discursivos. Estudos como os de Graves (2013), utilizando o banco de dados TIMIT, demonstraram o potencial das LSTM em superar outras abordagens em tarefas de transcrição automática, o que reforça a relevância de explorar esse modelo na área da acessibilidade.

Partindo dessa perspectiva, este projeto propõe investigar e implementar melhorias nos sistemas de geração automática de legendas baseados em IA, com foco na aplicação de redes LSTM. A expectativa é que, ao incorporar essas técnicas, seja possível alcançar maior precisão e sensibilidade linguística, promovendo uma experiência digital mais inclusiva para pessoas com deficiência auditiva e contribuindo para a democratização do acesso à informação em ambientes digitais.

2 MATERIAIS E MÉTODOS

Esta pesquisa será de natureza aplicada, com abordagem qualitativa e exploratória, voltada à identificação de falhas em modelos atuais de geração de legendas automáticas e à proposição de aprimoramentos com foco em acessibilidade para pessoas com deficiência auditiva.

A primeira etapa consiste em uma revisão bibliográfica sistemática, utilizando como fontes livros acadêmicos, artigos científicos indexados em bases como Scielo, IEEE Xplore e Google Scholar, além de documentos técnicos e legislações pertinentes, como a Lei Brasileira de Inclusão (Lei nº 13.146/2015). O objetivo desta etapa é consolidar os conceitos sobre acessibilidade digital, Processamento de Linguagem Natural (PLN), redes neurais recorrentes (RNNs), legendas automáticas e inclusão comunicacional.

Em seguida, será feita a coleta de dados empíricos por meio da observação não participante, selecionando uma amostra de 50 vídeos curtos com legendas geradas automaticamente por Inteligências Artificiais em plataformas de grande alcance como YouTube, TikTok e Instagram Reels. Os vídeos serão escolhidos com base em critérios como diversidade regional dos falantes, complexidade da fala e conteúdo educacional ou informativo.

Os dados coletados (legendas automáticas) serão analisados manualmente, buscando-se identificar e classificar os erros mais recorrentes, como falhas semânticas, truncamentos, trocas de palavras, problemas com sotaques ou expressões regionais. Essa análise será documentada em planilhas com categorias definidas para cada tipo de erro.

Com base nos resultados da análise, será realizado um estudo aprofundado sobre técnicas de PLN e redes neurais, em especial o modelo LSTM (Long Short-Term Memory), para projetar uma solução teórica de modelo de IA com maior sensibilidade à variação linguística e contextual da fala em português brasileiro. O modelo será proposto em nível de arquitetura, podendo futuramente ser implementado e testado computacionalmente em projetos futuros ou continuação da pesquisa.

Por fim, será elaborada uma avaliação teórica do impacto potencial do modelo proposto na acessibilidade de pessoas com deficiência auditiva, com base em estudos de caso e na comparação com os erros identificados nos modelos atuais. Os resultados serão consolidados em um relatório final, contendo análises, reflexões, limitações do estudo e recomendações para futuras pesquisas e aplicações práticas na área.



3 RESULTADOS E DISCUSSÕES

Espera-se que o desenvolvimento deste projeto de pesquisa contribua significativamente para o aprimoramento das tecnologias de geração automática de legendas, com foco na precisão semântica e na adaptação a variações linguísticas regionais. Como resultado, o estudo visa promover uma maior inclusão de pessoas com deficiência auditiva, por meio de soluções mais acessíveis, contextualizadas e alinhadas às diretrizes da Lei Brasileira de Inclusão. Além disso, espera-se gerar conhecimento técnico relevante nas áreas de Processamento de Linguagem Natural (PLN) e redes neurais recorrentes, fornecendo subsídios para futuros projetos de pesquisa e desenvolvimento de soluções inclusivas em plataformas educacionais e de mídia digital.

4 CONSIDERAÇÕES FINAIS

Investigar e propor melhorias na geração de legendas automáticas por meio de técnicas de Inteligência Artificial, com foco em acessibilidade, precisão contextual e inclusão de pessoas com deficiência auditiva.

Objetivos Específicos

- Analisar as limitações atuais de sistemas de Inteligência Artificial na transcrição de falas para legendas automáticas, especialmente em relação às variações linguísticas regionais.
- Estudar técnicas de Processamento de Linguagem Natural (PLN), com ênfase em Redes Neurais Recorrentes (RNNs) e suas variantes, para aprimorar a compreensão contextual da fala pelas IAs.
- Desenvolver um modelo experimental de geração de legendas com foco em precisão semântica, considerando sotaques, expressões idiomáticas e regionalismos da língua portuguesa.
- Avaliar a eficácia do modelo proposto por meio de testes comparativos com transcrições padrão, mensurando o ganho em acessibilidade e compreensão por usuários com deficiência auditiva.
- Verificar o potencial do modelo como ferramenta de apoio à acessibilidade digital em vídeos educativos, televisivos e em plataformas de streaming.

REFERÊNCIAS

FLECK, Leandro; TAVARES, Maria Hermínia Ferreira; EYNG, Eduardo; HELMANN, Andrieli Cristina; ANDRADE, Minéia Aparecida de Moares. **REDES NEURAIIS ARTIFICIAIS: PRINCÍPIOS BÁSICOS**. *Revista Eletrônica Científica Inovação e Tecnologia*, Medianeira, PR, v. 1, n. 13, p. 47-57, jan./jun. 2016.

AMAZON WEB SERVICES, INC. **O que é uma rede neural?**. Disponível em: <https://aws.amazon.com/what-is/neural-network/>. Acesso em: 29 abr. 2025.

AMAZON WEB SERVICES, INC. **O que é RNN?**. Disponível em: <https://aws.amazon.com/pt/what-is/recurrent-neural-network/>. Acesso em: 29 abr. 2025.

ARINOS, V., F.; FURLAN, K.; MACIEL, C.; SILVEIRA, M. S. **Legenda Automática do YouTube: essa inteligência artificial atende aos usuários?** In: *Anais do Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais*. SBC, 2019.



CASELI, Helena de Medeiros; NUNES, Maria das Graças Volpe (Org.). **Processamento de Linguagem Natural: Conceitos, Técnicas e Aplicações em Português**. 2. ed. São Carlos: BPLN, 2024.

LOPES, Carla; PERDIGÃO, Fernando. **Phoneme Recognition on the TIMIT Database**. In: IPSIC, Ivo (Ed.). *Speech Technologies*. InTech, 2011.

GRAVES, Alex; MOHAMED, Abdel-rahman; HINTON, Geoffrey. **SPEECH RECOGNITION WITH DEEP RECURRENT NEURAL NETWORKS**. arXiv:1303.5778v1 [cs.NE], 2013.