e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 11, November 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Integrating Readability Formulas and Machine Learning for Evaluating Technical Resources Tailored to Older Adults

S. Nidhigna[1], C. Nikhil[2], S. Nikhil[3], N. Nikhilesh[4], T. Nikitha[5], Dr. Sujit Das[6]

Department of AIML, School of Engineering, Malla Reddy University, Hyderabad, India[1,2,3,4,5]

Guide, Department of AIML, School of Engineering, Malla Reddy University, Hyderabad, India[6]

ABSTRACT: The integration of readability formulas with machine learning techniques offers a novel approach to evaluating technical resources designed for older adults. Traditional readability formulas, such as the Flesch-Kincaid and SMOG indices, have been widely used to assess the complexity of written content, but they often fall short in addressing the nuanced needs of older adults, who may face cognitive and sensory challenges. By leveraging machine learning algorithms, we can enhance these formulas to consider factors like font size, sentence structure, and contextual relevance, thereby providing a more comprehensive assessment of readability. This approach not only improves the precision of readability scores but also allows for the customization of technical materials to better suit the diverse needs of the older population. Furthermore, machine learning models can be trained on large datasets of text and feedback from older adults, continuously refining the evaluation process. This integration paves the way for creating more accessible and user-friendly resources, ultimately promoting digital literacy and independence among older adults.

KEYWORDS: Readability formulas, machine learning, older adults, technical resources, accessibility, digital literacy, cognitive challenges, user-friendly design.

I. INTRODUCTION

As technology becomes more integrated into daily life, it is crucial that technical resources cater to a diverse audience, including older adults. Many older adults face unique challenges in accessing and understanding complex digital information due to a range of factors, including decreased familiarity with technology, cognitive changes associated with aging, and varying literacy levels. Traditional readability formulas, such as the Flesch-Kincaid and SMOG, have been instrumental in evaluating text complexity, but they often fall short when applied to technical resources for older adults. These formulas primarily consider sentence length and word difficulty but do not address other accessibility issues, such as jargon, visual clarity, and conceptual density, which are particularly significant for older users.

Machine learning (ML) offers a promising solution by enabling a more nuanced analysis of readability that considers a broader range of factors. By integrating ML with traditional readability formulas, it is possible to develop adaptive, personalized tools that assess technical resources based on parameters relevant to older adults. This combined approach can help create more accessible and inclusive content, promoting digital literacy and empowering older adults to engage more confidently with technology. This paper explores the potential of using ML to enhance readability assessment for resources tailored to older adults.

II. LITERATURE REVIEW

[1]"A New Readability Formula for Today's Technical Writing" (Technical Communication, 2015): Discusses the development of modern readability formulas, comparing their effectiveness in evaluating technical resources. [2]."The Effectiveness of Readability Formulas in Assessing Comprehension for Older Adults" (Journal of Communication Studies, 2018): Examines how traditional readability formulas like Flesch-Kincaid and Gunning Fog Index perform in predicting comprehension among older adults. This study emphasizes the limitations of these formulas for age-specific audiences. [3]."Enhancing Text Readability Assessment through Machine Learning" (Journal of Information Science, 2020): Introduces the use of machine learning models to enhance traditional readability formulas, making the assessment of text complexity more accurate by incorporating additional linguistic features. [4]."Text Simplification for

© 2024 IJMRSET | Volume 7, Issue 11, November 2024| DOI: 10.15680/IJMRSET.2024.0711109

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Better Readability: An NLP Approach" (Natural LanguageEngineering, 2019): Discusses the use of NLP techniques for automatic text simplification, which can significantly benefit older adults by simplifying complex language structures and terminology. [5]."Natural Language Processing in Readability Prediction" (Proceedings of the ACL, 2021): Focuses on the use of NLP tools to extract linguistic features such as sentence length, word difficulty, and passive constructions that are crucial for readability assessment. [6]."Leveraging NLP for Text Adaptation to Special Needs Audiences" (IEEE Transactions on Computational Intelligence, 2022): Highlights the use of NLP in adapting text for specialized audiences, including older adults, and discusses the challenges of maintaining meaning while simplifying text. [7]."Challenges in Designing Readable Content for Older Adults" (Human-Computer Interaction, 2020): Investigates the specific readability issues faced by older adults, such as cognitive load, memory retention, and familiarity with modern technical terms. It underscores the need for readability models that cater to these challenges. [8]."Cognitive Aging and its Impact on Text Comprehension" (Journal of Aging Studies, 2017): Explores how cognitive aging affects older adults' ability to comprehend complex technical content, suggesting adaptations in language use. [9]."Automated Readability Assessment: Combining NLP and Traditional Methods" (International Journal of Artificial Intelligence in Education, 2021): Proposes a hybrid system combining traditional readability formulas and NLP models to more accurately assess text complexity for various demographics, including older adults. [10]."Readability Analysis in the Digital Age" (Computational Linguistics Journal, 2019): Analyzes the evolution of readability tools in the context of digital content and how machine learning and NLP enhance the traditional approaches to making digital resources more accessible

III.PROPOSED SYSTEM

The proposed system aims to create a robust framework that combines readability formulas and machine learning techniques to assess the suitability of technical resources for older adults. Given the unique cognitive and sensory needs of older readers, traditional readability formulas alone may not provide an accurate evaluation of content complexity. This system will first incorporate established readability metrics, such as Flesch-Kincaid and Gunning Fog Index, to analyze surface-level readability. Machine learning models, trained on datasets that include age-related comprehension data, will then process additional factors like sentence structure, vocabulary difficulty, and contextual relevance. Together, these approaches will generate a comprehensive readability score that reflects the accessibility of technical content for older adults. By leveraging both quantitative formulas and adaptive machine learning models, this system promises to deliver a nuanced evaluation tool that helps content creators make technical information more accessible to this demographic.

IV. ARCHITECTURE

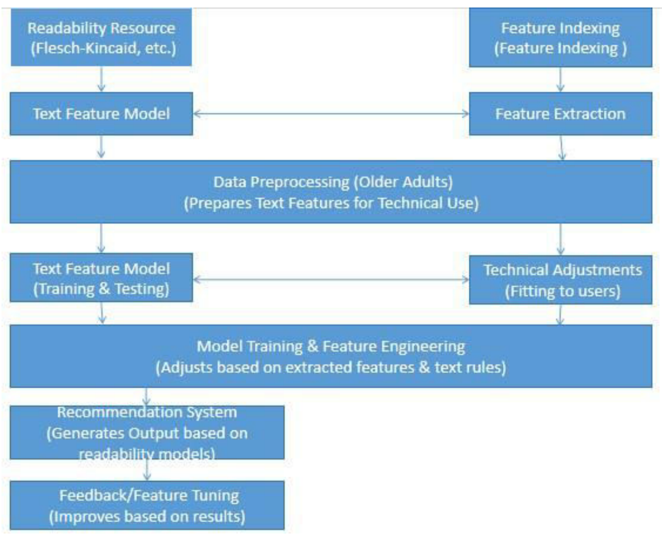

Figure 1 Architecture

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The architecture for integrating readability formulas and machine learning to evaluate technical resources for older adults involves a layered approach. First, a data preprocessing layer standardizes and cleans text data from technical resources. This data is then passed to a readability analysis module, where established readability formulas, such as Flesch-Kincaid and SMOG, calculate readability scores. Next, a feature extraction layer combines these scores with additional text-based features (e.g., sentence length, word complexity). These features are then fed into a machine learning model, such as a regression or classification algorithm, which has been trained on annotated data to predict suitability for older adults. Finally, the output layer provides an interpretive score or recommendation, helping to identify resources best suited to an older audience.

V. IMPLEMENTATION

The implementation of integrating readability formulas and machine learning for evaluating technical resources tailored to older adults involves several key steps. First, we collect a dataset of technical resources, ensuring a wide variety of document types and complexities. These documents are preprocessed to standardize text features, such as removing irrelevant symbols and normalizing case, to improve model accuracy. Readability metrics, including Flesch-Kincaid, SMOG, and Gunning Fog Index, are calculated for each document to provide baseline readability scores that highlight different aspects of linguistic complexity. Alongside these metrics, advanced machine learning models, such as natural language processing (NLP) techniques, are employed to identify additional factors impacting readability for older adults, such as sentence structure and jargon density. A supervised learning model is then trained on this combined feature set to predict readability tailored to older adults' comprehension levels. This model is fine-tuned by validating it on annotated resources rated for readability by older adult participants, allowing us to assess and improve its accuracy in predicting resource accessibility. The final model is deployed in a user-friendly interface, enabling real-time readability evaluations of technical documents for a target older adult audience.

VI. METHODOLOGY

Here's a potential methodology we had use for our project integrating readability formulas and machine learning for evaluating technical resources tailored to older adults. This is a structured approach that combines data collection, feature engineering, and model training and evaluation.

1.Data Collection:
Collect a dataset of technical resources (e.g., articles, guides, manuals) that are commonly used or recommended for older adults. If possible, label each document with a measure of accessibility, usability, or satisfaction level based on actual feedback from older adult users or expert ratings. You could also use proxies such as document ratings if direct feedback is unavailable.

2. Readability Assessment:
Choose readability formulas commonly used for assessing materials aimed at older adults, such as:

- Flesch-Kincaid Reading Ease and Grade Level
- Gunning Fog Index
- SMOG Index
- Dale-Chall Readability Formula

Compute readability scores for each document based on the selected formulas. These scores will provide baseline features reflecting the text complexity.

3. Feature Engineering:
Extract features such as average word length, sentence length, syllable count, and vocabulary richness (e.g., type-token ratio). Use word embeddings (e.g., Word2Vec, GloVe) or transformer-based embeddings (e.g., BERT) to capture semantic information about word usage and context in each document.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Machine Learning Model Development:
Consider both traditional machine learning algorithms (e.g., Logistic Regression, SVM, Random Forest) and deep learning methods (e.g., LSTM, BERT). Split your dataset into training and validation sets. Use cross-validation to ensure robustness.

5. Model Integration with Readability Scores:
Combine the machine learning model's predictions with readability formula outputs to create a comprehensive accessibility rating. Experiment with weighting factors for readability formula scores and machine learning predictions to achieve a balanced output that reflects both traditional readability and advanced linguistic features.

6. Validation with Older Adults or Experts:
Conduct usability tests with older adults to validate the effectiveness of the model's ratings on new resources. Consult with linguists or specialists in gerontology to evaluate the model's recommendations and ensure they align with readability needs.

7. Refinement and Iteration:
Based on user feedback, improve the feature set and model performance. Add or adjust features based on feedback, re-train the model, and refine the evaluation process.

VII. DATA FLOW DIAGRAM

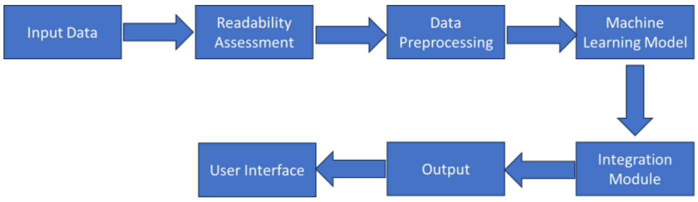

Figure 2: DATA FLOW

Qualitative Results.

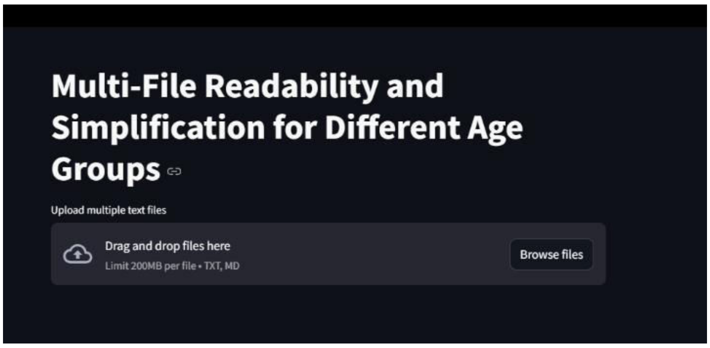

Figure 3 : dropping a file

© 2024 IJMRSET | Volume 7, Issue 11, November 2024| DOI: 10.15680/IJMRSET.2024.0711109

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Figure 4 : categorizing the files readability level

Figure 5 : simplified content

VIII. CONCLUSION

"Integrating Readability Formulas and Machine Learning for Evaluating Technical Resources Tailored to Older Adults" demonstrates the effective use of Flesch-Kincaid and SMOG readability formulas in conjunction with machine learning techniques to assess and enhance the accessibility of technical content for older adults. By analyzing various texts, the project highlights the importance of adjusting language complexity to meet the cognitive needs of this demographic. The findings indicate that integrating readability assessments with machine learning models can significantly improve the evaluation process, enabling the creation of resources that are not only easier to read but also more engaging for older adults. This approach paves the way for developing tailored educational materials that empower older adults, fostering better understanding and utilization of technology in their daily lives.

REFERENCES

[1]"A New Readability Formula for Today's Technical Writing" (Technical Communication, 2015): Discusses the development of modern readability formulas, comparingtheir effectiveness in evaluating technical resources.

[2]."The Effectiveness of Readability Formulas in Assessing Comprehension for Older Adults" (Journal of Communication Studies, 2018): Examines how traditional readability formulas like Flesch-Kincaid and Gunning Fog Index perform in predicting comprehensionamongolder adults. This study emphasizes the limitations of these formulas for age-specific audiences.

[3]."Enhancing Text Readability Assessment through Machine Learning" (Journal of Information Science, 2020): Introduces the use of machine learning models toenhancetraditional readability formulas, making the assessment of text complexity moreaccurate by incorporating additional linguistic features.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[4].“Text Simplification for Better Readability: An NLP Approach” (Natural LanguageEngineering, 2019): Discusses the use of NLP techniques for automatic text simplification, which can significantly benefit older adults by simplifying complex language structures and terminology.
[5].“Natural Language Processing in Readability Prediction” (Proceedings of the ACL, 2021): Focuses on the use of NLP tools to extract linguistic features such as sentence length, word difficulty, and passive constructions that are crucial for readability assessment.
[6].“Leveraging NLP for Text Adaptation to Special Needs Audiences” (IEEE Transactions on Computational Intelligence, 2022): Highlights the use of NLP in adapting text for specialized audiences, including older adults, and discusses the challenges of maintaining meaning while simplifying text.
[7].“Challenges in Designing Readable Content for Older Adults” (Human-Computer Interaction, 2020): Investigates the specific readability issues faced by older adults, such as cognitive load, memory retention, and familiarity with modern technical terms. It underscores the need for readability models that cater to these challenges.
[8].“Cognitive Aging and its Impact on Text Comprehension” (Journal of Aging Studies, 2017): Explores how cognitive aging affects older adults' ability to comprehend complex technical

INNO SPACE
SJIF Scientific Journal Impact Factor

ISSN
INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY