



Simple or Complex? Learning to Predict Readability of Bengali Texts



Susmoy Chakraborty^{1*}, Mir Tafseer Nayeem^{1*}, Wasi Uddin Ahmad²

¹Ahsanullah University of Science and Technology

²University of California, Los Angeles

***Equal Contribution, listed by alphabetical order**

SUMMARY

- Readability analysis of a low-resource language, **Bengali**
- English Readability analysis tool: Grammarly, Readable, etc. **No tool available** for the Bengali language
- Previous works: **Narrow** and sometimes **faulty**

Our Contributions

Age-to-age comparison to adapt English readability formulas, datasets, human-annotated corpora, dictionaries, algorithm, models, and... (Please see the next slide)

A Bengali Readability Analysis Tool!

BENGALI DOCUMENT READABILITY CHECKER

SIMPLE SENTENCE: GREEN, COMPLEX SENTENCE: RED

রাতের অন্ধকারে এক নেকড়ে চুকেছিল মানুষের গ্রামে। সেখানে কুকুরেরা তাকে ঘিরে এমন কামড়েছিল যে প্রাণ যাবার দশা হয়েছিল তার। কোনও রকমে প্রাণটা নিয়ে পালিয়ে আসতে পেরেছিল সে। কিন্তু কিছুদিনের মধ্যেই তার শরীরে কুকুরের কামড়ের ঘা বিধিয়ে উঠল। নেকড়ের হাটাচলার উপায় রইল না। যন্ত্রণায় কাতর হয়ে এক গাছতলায় গোঁজ হয়ে পড়ে রইল। **বিধিয়ে ওঠা ক্ষতের যন্ত্রণার ওপর ছিল পেটের টান।** বেচারি খিদেয় খুবই কাতর হয়ে পড়েছিল। এমন সময় সে দেখতে পেল, খানিক দূর দিয়ে একটা ভেড়া চলে যাচ্ছে। ছুটে গিয়ে শিকার ধরবার উপায় নেই। তাই সে কাতর গলায় ভেড়াকে ডেকে বলল, ও ভাই, শুনছো, একবারটি এদিকে এসো। ডাক শুনে ভেড়া দাঁড়িয়ে পড়ল। কিন্তু এগিয়ে না এসে বলল, কি বলতে চাও বল - আমি এখান থেকেই শুনতে পাব। নেকড়ে বলল, ভাই, ক্ষুধা পিপাসায় বড় কাতর হয়ে পড়েছি। তুমি যদি দয়া করে সামনের ঝরনা থেকে সামান্য জল এনে দাও, প্রাণটা বাঁচে। খাবারের ব্যবস্থা আমার কাছেই রয়েছে। শুনে ভেড়া বলল, ভাই, তোমার পিপাসার জল দিতে গিয়ে প্রাণটা দেওয়ার ইচ্ছে নেই। তুমি যে আমার ঘাড় ভেঙ্গেই তোমার খাবারের ব্যবস্থা করতে চাইছ, তা আমি বুঝতে পারছি। এই বলে সে সেখান থেকে দৌড়ে চলে গেল। নীতিকথা: ধূর্তের ছলনার অভাব হয়না, তাই মিষ্টি কথায় ভিজতে নেই। **ভবিষ্যতে সুখের আশা করে যারা বর্তমানে নিষ্কর্মা হয়ে বসে থাকে শেষ পর্যন্ত তাদের নিরাশাই হতে হয়।**

SUBMIT

INPUT DOCUMENT SUMMARY

READABILITY SCORE (OUT OF 100)	90.5
RATING	A
SENTENCE(S)	21
SIMPLE SENTENCE(S)	19
COMPLEX SENTENCE(S)	2
WORD(S)	203
AVERAGE WORDS PER SENTENCE	9.7
CONSONANT CONJUNCT(S)	30
ARI SCORE & AGE RANGE	5 & 10-11

CLEAR RESULTS

DOWNLOAD AS PDF

Thank You!



You can also mail us at

susmoyaust36@gmail.com || mir.nayeem@alumni.uleth.ca || wasiahmad@ucla.edu