|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Cat.** | **Method** | **Cat.** | **Method** | **Cat.** | **Method** |
| **RF** | Automated Readability Index (ARI) [44] | **WFF** | 25th percentil English Wikipedia | **HF** | # of Abbr tags |
| Coleman-Liau Index (CLI) [21] | 50th percentil English Wikipedia | # of A tags |
| Dale-Chall Index (DCI) [22] | 75th percentil English Wikipedia | # of Blockquote tags |
| Flesch-Kincaid Grade Level (FKGL) [23] | Mean Rank English Wikip. | # of Bold tags |
| Flesch Reading Ease (FRE) [23] | Mean Rank English Wikip. - Includes OV | # of Cite tags |
| Gunning Fog Index (GFI) [45] | 25th percentil Medical Reddit | # of Div tags |
| Lasbarhetsindex (LIX) [46] | 50th percentil Medical Reddit | # of Forms tags |
| Simple Measure of Gobbledygook (SMOG) [47] | 75th percentil Medical Reddit | # of H1 tags |
| **CRF** | # of Characters ⋆♦† | Mean Rank Medical Reddit | # of H2 tags |
| # of Words ⋆† | Mean Rank Medical Reddit - Includes OV | # of H3 tags |
| # of Sentences ⋆♦ | 25th percentil Pubmed | # of H4 tags |
| # of Difficult Words (Dale-Chall list [22]) ⋆♦† | 50th percentil Pubmed | # of H5 tags |
| # of Words Longer than 4 chars ⋆♦† | 75th percentil Pubmed | # of H6 tags |
| # of Words Longer than 6 chars ⋆♦† | Mean Rank Pubmed | #ofHs(any H above) |
| # of Words Longer than 10 chars ⋆♦† | Mean Rank Pubmed - Includes OV | # of Img tags |
| # of Words Longer than 13 chars ⋆♦† | 25th p. Wikipedia+Reddit+Pubmed | # of Input tags |
| # of Number of Syllables ⋆♦† | 50th p. Wikipedia+Reddit+Pubmed | # of Link tags |
| # of Polysyllable Words (>3 Syllables) ⋆♦† | Mean R. Wiki.+Reddit+Pubmed | # of DL tags |
| **NL** | # of Entities ⋆♦† | Mean R. Wiki.+Reddit+Pubmed - w. OV | # of Ul tags |
| # of verbs ⋆♦† | **GMV** | # of Words with Medical Prefix ⋆♦† | # of OL tags |
| # of nouns ⋆♦† | # of Words with Medical Suffix ⋆♦† | # of List (DL + UL + OL) |
| # of pronouns ⋆♦† | # of Acronyms ⋆♦† | # of Q tags |
| # of adjectives ⋆♦† | # of ICD Concepts ⋆♦† | # of Scripts tags |
| # of adverbs ⋆♦† | # of Drugbank ⋆♦† | # of Spans tags |
| # of adpositions ⋆♦† | # of Words in medical dict. (OpenMedSpel) ⋆♦ | # of Table tags |
| # of conjunctions ⋆♦† | **CMV** | CHV Mean Score for all Concepts ⋆♦† | # of P tags |
| # of determiners ⋆♦† | # of CHV Concepts ⋆♦† | **MLR** | Linear Regressor |
| # of cardinal numbers ⋆♦† | CHV Mean Score for Symptom Concepts ⋆♦† | Multi-layer Perceptron Regressor |
| # of particles or other function words ⋆♦† | # of CHV Symptom Concepts ⋆♦† | Random Forest Regressor |
| # of other POS (foreign words, typos) ⋆♦† | CHV Mean Score for Disease Concepts ⋆♦† | Support Vector Machine Regressor |
| # of punctuation ⋆♦† | # of CHV Disease Concepts ⋆♦† | Gradient Boosting Regressor |
| Height of part-of-speech parser tree ⋆♦† | **EMV** | # of MeSH Concepts ⋆♦† | Logistic Regression |
| # of Stopwords ⋆♦† | Average Tree of MeSH Concepts ⋆♦† | **MLC** | Multi-layer Perceptron Classifier |
| # of words not found in Aspell Eng. dict. ⋆♦† | # of MeSH Symptom Concepts ⋆♦† | Random Forest Classifier |
| Positive Words ⋆♦† | Average Tree of MeSH Symptom Concepts ⋆♦† | Support Vector Machine Classifier |
| Negative Words ⋆♦† | # of MeSH Disease Concepts ⋆♦† | Multinomial Naive Bayes |
| Neural Words ⋆♦† | Average Tree of MeSH Disease Concepts ⋆♦† | Gradient Boosting Classifier |