Passos:

Eliminant guions i caracters com “ < “ . Substituint “n’t” o “‘ ’ m” per “not” i “am”

ref: <https://gebakx.github.io/ihlt/sts/docs/2-takelab.pdf> <- apartat preprocessing

Eliminem stopwords