Notas:

* O erro ainda está alto quando usadas todas as classes (~0,53).
* Antes de remover as 100 palavras mais freqüentes, a classificação errava muito mais (~0,83)