V1: Pure word vectorisation with logistic regression.

Tf-Idf (bag of words):

C=10, 1:1 class weighting

A screen shot of a computer

Description automatically generated

C=10, 1:100 class weighting

A screen shot of a computer

Description automatically generated

C=0.1, 1:100 class weighting

A screenshot of a computer screen

Description automatically generated

Word2Vec:

C=0.1, 1:100 class weighting

A screenshot of a computer

Description automatically generated

V2: Logistic regression with features:

N authors

Year published/ age

Title length

Abstract length

Abstract length trimmed

Keyword based features

Dm

Hb!

I hope you have a nice day and take a break from worrying about global affairs!

C=0.1, 1:60 class weighting

A screenshot of a computer

Description automatically generated