Yohanes Dimas Pratama

A11.2021.13254 - A11.4705

Klasifikasi Dokumen dengan Naive Bayes

1. Melakukan tokenisasi dan stemming pada D7

D7 = ”Demo buruh menaikkan upah minimum”



1. Menghitung probabilitas prior

* Jumlah seluruh dokumen (N) = 6
* Jumlah dokumen dalam kategori Olahraga (N\_Olahraga) = 2
* Jumlah dokumen dalam kategori Politik (N\_Politik) = 2
* Jumlah dokumen dalam kategori Ekonomi (N\_Ekonomi) = 2

Masukkan ke rumus

* P(Olahraga) =
* P(Politik) =
* P(Ekonomi) =

1. Menghitung probabilitas bayes

* Olahraga:
* Politik:
* Ekonomi:

1. Normalisasi kategori menggunakan rumus

Normalisasi probabilitas:

* Olahraga: *= 0.37*
* Politik: = 0.23
* Ekonomi: = 0.40

1. Kesimpulan

Karena nilai probabilitas terbesar adalah kategori Ekonomi maka dapat disimpulkan bahwa D7 termasuk ke dalam kategori Ekonomi.