



Instituto Politécnico Nacional
Escuela Superior de Cómputo

Polaridad y sentimientos

Natural Language Processing

Estudiante:

Nicolás Sayago Abigail

Profesora:

Olga Kolesnicova

Mayo 26, 2020

1 Resultados

Estos son los resultados generales, es decir, la polaridad obtenida por cada rank. Podemos observar que como era de esperarse, las películas que recibieron una estrella, obtuvieron la polaridad más baja, también se observa que en el rank 3 y 4 la polaridad es muy parecida. Finalmente aunque se esperaría que la mejor polaridad la tuviera el rank 5, eso no necesariamente pasa.

```
-----> RANK: 1   # Reviews: 351   Total Polarity: 0.12774928774928776
-----> RANK: 2   # Reviews: 923   Total Polarity: 0.16586132177681473
-----> RANK: 3   # Reviews: 1253  Total Polarity: 0.20695929768555466
-----> RANK: 4   # Reviews: 890   Total Polarity: 0.20946067415730338
-----> RANK: 5   # Reviews: 461   Total Polarity: 0.19830802603036876
```

Estos son algunos ejemplos de los valores de polaridad de cada rank.

RANK 1: Se observan en su mayoría negativas.

```
-----> RANK: 1   # Reviews: 351   Total Polarity: 0.12774928774928776
[0.24, 0.06, 0.23, 0.27, 0.18, 0.03, -0.05, 0.13, -0.02, 0.15, 0.05, 0.16, 0.14, 0.01, 0.17, -0.01, 0.17,
0.17, -0.02, 0.05, 0.28, 0.23, 0.22, 0.34, 0.28, 0.23, 0.19, 0.24, 0.07, 0.29, 0.15, -0.1, 0.27, 0.2, 0.18,
0.13, 0.12, 0.09, 0.08, 0.27, 0.21, 0.38, 0.27, 0.04, 0.08, 0.09, 0.16, 0.03, -0.2, 0.12, 0.19, 0.2, 0.26,
-0.04, 0.38, -0.05, 0.08, 0.14, 0.11, -0.07, 0.05, 0.03, 0.12, 0.17, 0.29, 0.02, -0.1, 0.28, -0.0, 0.05, 0.34,
0.24, 0.18, 0.02, 0.15, 0.02, 0.25, 0.12, 0.16, 0.01, 0.15, 0.14, 0.17, 0.26, 0.3, -0.01, 0.16, -0.04, 0.24,
0.21, 0.2, 0.13, 0.04, 0.11, -0.0, 0.21, 0.18, 0.26, 0.23, 0.07, 0.13, 0.17, 0.13, 0.09, 0.24, 0.1, 0.42,
0.38, -0.09, -0.17, 0.34, 0.27, 0.04, 0.19, 0.22, 0.05, 0.12, 0.29, 0.22, 0.14, 0.01, 0.02, 0.16, 0.1, 0.17,
0.11, 0.12, 0.14, 0.23, 0.18, -0.18, 0.26, 0.18, 0.17, 0.31, -0.17, 0.21, 0.23, 0.15, 0.13, 0.09, 0.0, 0.09,
0.32, 0.09, 0.12, 0.16, 0.29, -0.06, 0.1, 0.04, -0.01, 0.14, 0.05, 0.07, 0.18, 0.01, -0.08, 0.07, 0.03, 0.08,
0.08, 0.14, -0.13, 0.16, 0.22, 0.1, 0.07, 0.1, -0.13, -0.09, 0.15, 0.09, 0.19, -0.03, 0.06, 0.16, 0.12, 0.13,
0.1, 0.3, 0.09, 0.12, 0.12, 0.04, 0.24, 0.07, -0.08, 0.11, 0.06, 0.25, 0.12, 0.16, 0.08, 0.1, 0.25, -0.11,
0.31, 0.12, 0.03, 0.15, 0.01, 0.15, 0.17, 0.13, 0.27, 0.13, 0.24, 0.31, 0.27, 0.21, 0.17, 0.13, 0.24, 0.32,
0.15, 0.08, 0.17, 0.22, 0.35, 0.1, 0.17, 0.08, 0.11, 0.2, 0.11, 0.1, 0.21, 0.21, 0.26, 0.12, 0.07, 0.29,
-0.06, 0.14, 0.07, 0.11, 0.07, 0.12, 0.23, -0.05, 0.25, 0.16, 0.06, 0.22, 0.08, 0.27, 0.18, -0.0, -0.07, 0.26,
0.01, 0.28, 0.14, 0.3, 0.1, 0.26, -0.04, 0.09, 0.23, 0.23, 0.28, -0.03, -0.0, 0.25, 0.09, 0.01, 0.12, 0.34,
0.23, 0.07, 0.11, 0.2, -0.08, 0.04, 0.17, 0.15, 0.2, -0.02, 0.16, -0.18, 0.03, 0.15, 0.21, 0.05, 0.01, 0.16,
0.08, 0.42, 0.04, -0.03, -0.05, 0.24, 0.13, -0.08, 0.04, 0.25, 0.06, -0.03, 0.03, 0.11, 0.13, 0.17, 0.04, 0.1,
0.21, 0.26, -0.13, 0.27, 0.05, -0.11, -0.01, 0.03, 0.06, 0.06, 0.17, 0.12, 0.29, 0.13, 0.08, 0.1, 0.18, 0.17,
0.19, -0.05, 0.26, 0.02, 0.11, 0.14, 0.04, 0.11, 0.12, 0.09, 0.06, 0.14, 0.21, 0.18, 0.07, 0.27, -0.1, 0.31,
0.19, 0.19, 0.23, 0.21, 0.16, 0.18, 0.08, 0.25, 0.09, 0.13]
```

RANK 2

```

-----> RANK: 2   # Reviews: 923   Total Polarity: 0.16586132177681473
[-0.01, 0.21, 0.15, 0.37, 0.24, 0.28, 0.06, 0.18, 0.13, 0.19, -0.07, 0.04, 0.14, 0.11, 0.26, -0.02, 0.16,
0.12, 0.17, 0.15, 0.15, 0.16, 0.28, 0.31, 0.03, 0.15, 0.18, 0.13, 0.25, 0.07, 0.09, 0.43, 0.13, 0.31, 0.19,
0.14, 0.15, 0.19, 0.17, 0.26, 0.22, 0.2, -0.2, 0.03, 0.22, 0.11, 0.04, 0.13, 0.15, 0.16, 0.01, 0.1, 0.14, 0.0,
0.27, 0.09, 0.16, 0.09, 0.17, 0.06, 0.33, 0.23, 0.21, 0.18, 0.3, 0.19, 0.09, 0.2, 0.15, 0.36, 0.32, 0.27,
0.27, 0.3, 0.28, 0.15, 0.03, 0.16, 0.14, 0.25, 0.2, 0.18, 0.18, 0.16, -0.06, 0.32, 0.26, 0.29, 0.18, 0.13,
0.31, 0.34, 0.13, 0.22, 0.14, 0.18, 0.33, 0.27, 0.27, 0.08, 0.03, 0.04, 0.35, 0.18, 0.27, 0.14, 0.17, 0.23,
0.2, 0.19, 0.26, 0.36, 0.19, 0.13, 0.13, 0.18, 0.16, 0.28, 0.21, 0.06, 0.26, 0.18, 0.2, 0.24, 0.12, 0.23,
0.25, 0.17, 0.27, 0.28, 0.18, 0.14, 0.07, 0.17, 0.22, 0.22, 0.09, 0.08, 0.13, 0.02, 0.21, 0.2, 0.04, 0.13,
0.02, 0.26, 0.23, 0.05, -0.04, 0.31, -0.01, 0.17, 0.15, 0.08, 0.34, 0.25, -0.09, 0.06, 0.17, 0.09, 0.12, 0.18,
0.2, 0.09, 0.22, 0.13, 0.16, 0.23, 0.28, 0.17, 0.22, 0.33, 0.17, 0.11, 0.07, 0.27, 0.2, 0.28, 0.22, 0.1, 0.28,
0.12, 0.15, 0.22, 0.13, 0.09, 0.31, 0.12, 0.27, -0.15, 0.13, 0.34, 0.26, 0.09, 0.04, 0.17, 0.24, 0.11, 0.28,
0.26, 0.15, -0.0, 0.07, 0.29, 0.15, 0.22, 0.02, 0.34, 0.2, 0.11, -0.0, 0.04, 0.09, 0.17, 0.11, 0.14, 0.11,
0.13, 0.33, 0.17, 0.24, 0.15, 0.21, 0.15, 0.21, 0.33, 0.29, 0.4, 0.06, 0.09, 0.34, 0.23, 0.28, 0.2, 0.16,
0.04, 0.23, -0.2, 0.07, 0.34, 0.1, 0.21, 0.21, 0.11, 0.15, 0.4, 0.19, 0.22, 0.26, 0.12, 0.21, 0.28, 0.33,
0.24, 0.21, 0.24, -0.06, 0.21, 0.06, 0.14, 0.19, 0.23, 0.07, 0.21, 0.27, 0.14, 0.34, 0.19, 0.11, 0.12, 0.21,
0.02, 0.11, 0.13, 0.1, 0.13, -0.36, 0.25, 0.1, 0.25, 0.25, 0.09, 0.11, 0.11, 0.18, 0.28, 0.06, 0.28, 0.14,
0.17, -0.01, 0.29, 0.25, 0.28, 0.26, 0.19, 0.17, 0.19, 0.19, 0.16, 0.14, 0.47, 0.24, 0.01, 0.12, 0.08, 0.21,
0.27, 0.1, 0.15, 0.01, 0.01, 0.33, 0.44, 0.09, 0.26, 0.15, 0.29, -0.05, 0.24, -0.0, 0.18, 0.16, 0.09, 0.08,
0.08, -0.08, -0.01, 0.11, 0.23, 0.16, 0.04, 0.35, 0.52, 0.12, 0.3, 0.07, 0.25, 0.16, 0.24, 0.1, 0.05, 0.06,
0.16, 0.04, 0.14, 0.31, 0.21, 0.18, 0.12, 0.16, 0.18, 0.17, 0.27, 0.1, 0.12, 0.24, 0.34, 0.23, 0.23, 0.34,
0.08, 0.14, -0.18, 0.27, -0.04, 0.37, 0.16, 0.27, -0.07, 0.05, 0.08, 0.19, 0.09, 0.16, 0.11, 0.17, 0.09, 0.06,
0.08, 0.19, 0.25, 0.09, 0.21, 0.3, 0.15, 0.16, 0.19, 0.1, 0.25, 0.36, 0.21, 0.07, 0.0, 0.15, 0.13, 0.11, 0.24,
0.06, 0.12, 0.24, 0.2, 0.13, 0.15, 0.25, 0.16, 0.26, 0.14, 0.1, 0.22, 0.25, -0.03, 0.07, 0.15, 0.17, 0.25,
0.19, 0.37, 0.27, -0.05, 0.16, 0.27, 0.28, 0.12, 0.1, 0.28, 0.06, 0.21, 0.32, 0.06, 0.27, 0.16, 0.21, 0.12,
0.19, 0.32, 0.04, 0.09, -0.17, 0.27, 0.02, 0.24, 0.32, 0.15, 0.2, 0.25, 0.08, 0.16, 0.09, 0.09, 0.01, 0.1,

```

RANK 3

```

-----> RANK: 3   # Reviews: 1253   Total Polarity: 0.20695929768555466
[0.17, 0.11, 0.38, 0.14, 0.3, 0.0, 0.01, 0.25, 0.06, 0.32, 0.26, 0.07, 0.07, 0.04, 0.14, 0.21, 0.11, 0.19,
0.01, 0.21, 0.14, 0.16, 0.08, -0.23, 0.22, 0.12, 0.15, 0.22, 0.29, 0.14, 0.26, -0.04, 0.18, 0.42, 0.04, 0.25,
0.25, 0.29, -0.1, 0.23, 0.36, 0.2, 0.09, 0.17, 0.3, 0.1, 0.32, -0.01, 0.25, 0.43, 0.33, 0.02, 0.18, 0.25,
0.19, 0.22, 0.21, 0.34, 0.2, 0.13, 0.11, 0.32, 0.24, 0.13, 0.2, 0.17, 0.27, 0.24, 0.24, 0.23, 0.35, 0.13,
0.27, 0.14, 0.06, 0.09, 0.13, 0.1, 0.14, 0.36, 0.36, 0.13, 0.29, 0.1, 0.22, 0.26, 0.19, 0.15, 0.27, 0.11,
0.27, -0.0, 0.31, 0.17, -0.01, 0.28, 0.18, 0.22, 0.2, 0.38, 0.25, -0.12, 0.3, 0.29, 0.3, 0.34, 0.24, 0.22,
0.14, 0.24, 0.23, 0.29, 0.16, 0.34, 0.19, 0.16, 0.23, 0.25, 0.19, 0.22, 0.25, 0.26, 0.16, 0.31, 0.38, 0.45,
0.36, 0.21, 0.02, 0.21, -0.09, 0.31, 0.3, -0.02, 0.26, 0.26, 0.22, 0.33, 0.0, 0.32, -0.05, 0.18, 0.21, 0.17,
0.24, 0.28, 0.14, -0.24, 0.3, 0.36, 0.27, 0.39, 0.14, 0.29, 0.1, 0.15, 0.29, 0.36, 0.27, 0.17, 0.36, 0.37,
0.19, 0.28, 0.32, 0.19, 0.31, 0.28, 0.22, 0.18, 0.0, 0.2, 0.18, 0.27, 0.19, 0.22, 0.21, 0.26, 0.26, 0.3, 0.34,
0.13, 0.23, 0.15, 0.19, 0.1, 0.3, 0.27, 0.13, -0.01, 0.16, 0.09, 0.25, 0.16, 0.23, 0.11, 0.13, 0.21, 0.1,
0.26, 0.09, 0.26, 0.19, 0.02, 0.0, 0.13, 0.27, 0.2, 0.17, 0.01, 0.28, 0.28, 0.31, 0.26, -0.05, -0.09, 0.38,
0.16, 0.17, 0.14, 0.13, 0.25, 0.23, 0.1, 0.13, 0.46, 0.2, 0.18, 0.17, 0.32, 0.23, 0.24, 0.11, 0.36, 0.19,
0.27, 0.12, 0.01, 0.31, 0.21, 0.14, 0.2, -0.09, 0.25, 0.07, 0.11, -0.01, 0.21, 0.12, -0.03, 0.28, 0.09, 0.16,
0.09, 0.26, -0.05, 0.16, 0.08, 0.09, 0.22, 0.13, 0.3, 0.28, 0.05, 0.01, 0.18, 0.02, 0.19, 0.36, 0.21, 0.27,
0.27, 0.2, 0.29, 0.38, 0.18, 0.5, 0.29, 0.26, 0.17, 0.31, 0.22, 0.48, 0.09, 0.26, -0.04, 0.11, 0.33, 0.15,
0.3, 0.13, 0.31, 0.17, 0.14, 0.2, 0.26, 0.26, 0.16, 0.08, 0.12, 0.35, 0.3, 0.33, 0.29, 0.07, 0.22, -0.09,
-0.08, 0.35, 0.31, 0.18, 0.14, 0.12, 0.14, 0.3, 0.59, 0.03, 0.14, 0.24, 0.28, 0.32, 0.1, 0.14, 0.24, 0.17,
0.32, 0.09, 0.23, 0.36, 0.47, 0.42, 0.17, 0.33, 0.03, 0.06, 0.19, 0.26, 0.26, 0.36, 0.09, -0.01, 0.34, 0.02,
0.38, 0.2, 0.07, 0.3, 0.17, 0.19, 0.35, 0.22, 0.13, 0.32, 0.18, 0.21, 0.06, 0.14, 0.15, 0.24, 0.2, 0.16, 0.22,
0.47, 0.01, 0.19, 0.28, 0.16, 0.15, 0.17, 0.11, 0.2, 0.27, 0.23, 0.44, 0.26, 0.13, 0.21, 0.33, 0.02, 0.17,

```

RANK 4

```

-----> RANK: 4 # Reviews: 890 Total Polarity: 0.20946067415730338
[0.04, 0.21, 0.35, 0.16, 0.03, 0.28, 0.26, 0.34, 0.15, 0.11, 0.11, 0.05, 0.4, 0.22, 0.24, 0.19, 0.0, 0.34,
0.17, 0.18, 0.17, 0.08, 0.1, 0.05, 0.26, 0.34, 0.36, 0.17, 0.25, 0.34, 0.19, 0.2, 0.28, 0.22, 0.22, 0.1, 0.11,
0.09, 0.21, 0.22, 0.15, 0.17, 0.24, 0.09, 0.34, 0.31, 0.2, 0.09, 0.14, 0.02, 0.11, 0.29, 0.24, 0.11, 0.37,
0.13, 0.28, 0.14, 0.14, 0.17, 0.02, 0.27, 0.23, 0.16, 0.13, 0.36, 0.31, 0.21, 0.16, 0.14, 0.14, 0.33, 0.28,
0.29, 0.25, 0.04, 0.35, 0.23, 0.21, 0.23, 0.25, 0.39, 0.29, 0.16, 0.12, 0.0, 0.3, 0.07, 0.24, 0.04, 0.19, 0.1,
0.15, 0.33, 0.25, 0.17, 0.4, 0.16, 0.4, -0.01, 0.36, 0.41, -0.02, 0.27, 0.07, 0.33, 0.25, 0.21, 0.3, -0.02,
0.23, -0.01, 0.26, -0.01, 0.36, 0.16, 0.22, 0.1, 0.29, 0.02, 0.09, 0.15, 0.34, 0.41, 0.27, 0.2, 0.29, 0.25,
-0.01, 0.24, 0.22, 0.29, -0.1, 0.28, 0.17, 0.22, 0.13, -0.11, 0.03, 0.05, 0.29, 0.2, 0.09, 0.33, 0.49, 0.2,
0.01, 0.27, 0.18, 0.21, 0.43, 0.1, 0.16, 0.21, 0.2, 0.23, 0.12, 0.34, 0.36, 0.23, 0.19, 0.32, 0.27, 0.23,
0.29, 0.13, 0.18, 0.23, 0.29, 0.16, 0.03, 0.14, 0.21, 0.2, 0.2, 0.09, 0.34, 0.26, 0.25, 0.3, 0.31, 0.18, 0.18,
0.07, 0.16, 0.09, 0.13, 0.36, 0.21, 0.21, 0.2, 0.04, 0.39, 0.21, 0.21, 0.3, 0.13, 0.2, 0.42, 0.23, 0.09, 0.22,
0.1, 0.29, 0.23, 0.16, 0.14, 0.29, 0.28, 0.34, 0.37, 0.16, 0.32, 0.4, 0.33, 0.35, -0.01, 0.13, 0.21, 0.16,
0.34, 0.1, 0.33, 0.17, 0.02, 0.09, 0.08, 0.26, 0.15, 0.24, 0.26, 0.38, 0.12, 0.18, 0.05, 0.28, 0.14, 0.33,
0.3, 0.25, 0.23, 0.07, 0.18, -0.03, 0.11, 0.18, 0.24, 0.25, 0.21, 0.2, 0.16, 0.17, 0.1, 0.21, 0.24, 0.18,
0.19, 0.1, -0.03, 0.23, 0.22, 0.34, 0.29, 0.37, 0.1, 0.23, 0.07, 0.17, 0.24, 0.1, 0.08, 0.38, 0.34, 0.18,
-0.0, 0.02, 0.4, 0.3, 0.19, 0.14, 0.1, 0.11, 0.19, 0.31, 0.23, 0.17, 0.15, 0.32, 0.23, 0.34, 0.22, 0.28, 0.11,
0.19, 0.25, 0.05, 0.18, 0.29, 0.04, 0.13, 0.2, 0.21, 0.07, 0.22, 0.28, 0.2, 0.26, 0.15, 0.29, 0.06, 0.17,
0.26, 0.01, 0.22, 0.25, 0.26, 0.38, 0.4, 0.18, 0.0, 0.1, 0.25, 0.28, 0.11, 0.13, 0.28, 0.23, 0.34, 0.26, 0.2,
0.27, 0.27, 0.15, 0.12, 0.23, 0.25, 0.06, 0.14, 0.28, 0.19, 0.25, 0.17, 0.26, 0.06, 0.23, 0.09, 0.14, 0.32,
0.26, 0.25, 0.2, 0.11, 0.31, 0.31, 0.18, 0.22, 0.28, 0.3, 0.3, 0.17, 0.37, 0.17, -0.0, 0.3, 0.27, 0.2, 0.26,
0.17, 0.16, 0.34, 0.25, 0.35, 0.18, 0.47, 0.22, 0.35, 0.07, 0.29, 0.32, 0.22, 0.22, 0.33, 0.29, 0.19, 0.4,
0.23, 0.46, 0.09, 0.17, 0.29, 0.2, 0.28, 0.29, 0.25, 0.23, 0.34, 0.21, 0.32, -0.04, 0.26, 0.17, 0.33, 0.33,
0.15, 0.41, 0.19, 0.14, 0.22, 0.21, 0.2, 0.18, 0.3, 0.22, 0.31, 0.23, 0.17, 0.26, 0.16, 0.15, 0.19, 0.12,
0.17, 0.17, 0.13, 0.35, 0.15, 0.22, 0.04, 0.08, 0.2, 0.24, 0.29, 0.16, 0.27, 0.18, 0.26, 0.02, 0.23, 0.34,

```

RANK 5

```

-----> RANK: 5 # Reviews: 461 Total Polarity: 0.19830802603036876
[0.1, 0.15, 0.33, 0.27, 0.02, 0.11, 0.17, 0.09, -0.01, 0.07, 0.01, 0.24, 0.34, 0.05, 0.01, 0.14, 0.1, 0.14,
0.12, 0.15, -0.0, -0.05, 0.18, 0.14, 0.26, 0.2, 0.23, 0.29, 0.2, 0.32, 0.05, 0.22, 0.05, -0.11, 0.17, 0.12,
0.07, 0.29, 0.15, 0.31, 0.09, 0.16, 0.15, 0.32, 0.38, 0.31, -0.14, 0.16, 0.2, 0.22, 0.31, 0.41, 0.14, 0.15,
0.0, 0.11, 0.05, 0.19, 0.19, 0.32, 0.17, 0.29, -0.14, 0.18, 0.4, 0.1, 0.1, 0.12, 0.11, 0.27, 0.16, 0.34, 0.15,
0.33, 0.21, 0.24, 0.07, 0.35, 0.1, 0.22, 0.29, 0.25, 0.31, 0.19, 0.17, 0.1, 0.28, 0.29, 0.22, 0.18, 0.16, 0.2,
0.29, 0.26, 0.15, 0.31, 0.18, 0.07, 0.1, 0.32, 0.29, 0.06, 0.25, 0.14, -0.04, 0.12, 0.21, 0.43, 0.32, 0.34,
0.22, 0.13, 0.31, 0.16, 0.08, 0.22, 0.21, 0.1, 0.16, 0.18, 0.21, 0.3, 0.08, 0.22, 0.3, 0.22, 0.26, 0.39, 0.12,
0.28, 0.16, 0.18, 0.19, -0.09, 0.11, 0.21, 0.27, 0.24, 0.37, 0.25, 0.34, 0.35, 0.15, 0.42, 0.38, 0.17, 0.35,
0.24, -0.0, 0.2, 0.16, 0.05, 0.17, 0.17, 0.04, 0.11, 0.11, 0.13, 0.06, 0.1, 0.09, 0.29, 0.21, 0.16, 0.25,
0.08, 0.27, 0.18, 0.16, 0.39, -0.02, 0.27, 0.22, 0.23, 0.16, 0.18, 0.21, 0.32, 0.14, -0.04, 0.17, 0.26, 0.27,
0.16, 0.05, 0.24, 0.37, 0.32, 0.09, 0.19, 0.19, 0.27, 0.13, 0.38, 0.38, 0.18, 0.13, 0.32, 0.18, 0.13, 0.13,
0.23, 0.17, 0.18, 0.15, 0.11, 0.14, 0.26, 0.15, 0.34, 0.24, -0.02, 0.11, 0.2, 0.11, 0.38, 0.1, 0.2, 0.26,
0.15, 0.32, 0.32, 0.33, 0.22, 0.31, 0.15, 0.23, 0.32, 0.28, 0.3, 0.35, 0.22, 0.3, 0.13, 0.28, 0.22, 0.25,
0.33, 0.22, 0.23, 0.15, 0.17, 0.39, 0.22, 0.17, 0.31, 0.05, 0.31, 0.27, 0.14, 0.33, 0.28, 0.3, 0.25, 0.07,
0.39, 0.26, -0.01, 0.22, 0.33, 0.3, 0.24, 0.22, 0.0, 0.19, 0.32, 0.07, 0.3, 0.35, 0.23, 0.29, 0.4, 0.25, 0.28,
0.12, 0.19, 0.26, 0.23, 0.2, 0.32, 0.06, 0.06, 0.1, 0.15, -0.09, 0.32, 0.22, 0.03, 0.2, 0.19, 0.28, 0.04,
0.33, 0.18, 0.29, 0.18, 0.06, 0.45, 0.24, 0.02, 0.28, 0.34, 0.34, 0.03, 0.4, 0.2, 0.33, 0.16, 0.14, 0.25,
0.33, 0.33, 0.08, 0.25, 0.3, 0.36, 0.37, -0.03, 0.18, 0.44, 0.07, 0.11, 0.28, 0.19, 0.37, 0.38, 0.13, 0.19,
0.21, 0.16, -0.12, 0.2, 0.19, 0.1, 0.3, 0.22, 0.17, 0.17, 0.06, 0.07, 0.12, 0.17, 0.22, 0.21, 0.07, 0.2, 0.34,
0.26, 0.48, 0.12, 0.2, 0.28, 0.11, 0.07, 0.21, 0.1, 0.24, 0.28, 0.16, 0.13, 0.37, 0.42, 0.22, 0.18, 0.25,
-0.09, 0.22, 0.38, 0.35, 0.2, 0.06, 0.05, 0.15, 0.05, 0.2, -0.0, 0.17, 0.29, 0.35, 0.35, 0.12, 0.18, 0.28,
0.22, 0.11, 0.21, 0.03, 0.14, 0.19, 0.09, 0.22, 0.25, 0.17, 0.35, 0.27, 0.08, 0.08, 0.37, 0.05, 0.26, 0.3,
0.17, 0.16, 0.26, 0.23, 0.24, 0.25, 0.33, 0.16, 0.15, 0.04, 0.06, -0.02, 0.18, 0.35, 0.19, 0.43, 0.26, 0.14,

```


✓ Código fuente

```
import nltk
import re
import math
import numpy as np
import mord
from bs4 import BeautifulSoup
from sklearn.feature_extraction.text import TfidfVectorizer
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
from sklearn import metrics

#####
#                               GETTING DICTIONARY
#####
def getWord(s):
    word = ''
    j = s.find('">')
    j = j + 3
    while(s[j] != '<'):
        word += s[j]
        j = j + 1
    word = word[:len(word)-1]
    return word

# Return a string
def getPOS(s):
    j = s.find('pos=')
    j = j + 5
    tag = ''
    while(s[j] != '"'):
        tag = tag + s[j]
        j = j + 1
    return tag

# Return a number that represent polarity
def getPolarity(s):
    j = s.find('pol=')
    j = j + 5
    pol = ''
    while(s[j] != '"'):
        pol = pol + s[j]
        j = j + 1
```

```

    pol = float(pol)
    return pol

# Return dictionary of pairs with polarity
def getDictionary(corpusRoot, code):
    dictionary = {}
    f = open(corpusRoot, encoding = code)
    text = f.readlines()
    f.close()

    for i in range(0, len(text)):
        word = getWord(text[i])
        POS = getPOS(text[i])
        polarity = getPolarity(text[i])
        pair = (word, POS)
        dictionary[pair] = polarity
    return dictionary

#####
#                               GETTING CORPUS TEXT
#####
# Return vector of vector of tuples that represent reviews
def getText(corpusRoot, code, n):
    reviews = list()
    for i in range(2, n):
        try:
            f = open(corpusRoot + str(i) + ".review.pos",
                    encoding = code) #Cod: utf-8, latin-1
            text = f.readlines()
            tokens = [nltk.word_tokenize(line) for line in text]
            review = list()
            for line in tokens:
                if len(line) > 0:
                    tag = 'n'
                    if len(line) > 3:
                        tag = line[2][0].lower()
                    review.append((line[1], tag))
            reviews.append(review)
            f.close()
        except:
            continue
    return reviews

# Return vector of ranks

```

```

def getRank(corpusRoot, code, n):
    ranks = list()
    for i in range(2, n):
        try:
            f = open(corpusRoot + str(i) + ".xml")
            lines = f.readlines()
            j = lines[0].index( ' rank=' )
            ranks.append(int(lines[0][j+7]))
            f.close()
        except:
            continue
    return ranks

def removeStopwords(tokens, language):
    sw = stopwords.words(language)
    clean = list()
    for review in tokens:
        aux = list()
        for word in review:
            if word[0] not in sw:
                aux.append(word)
        clean.append(aux)
    return clean

#####
# GET RESULT
#####
def getMatrix(dictionary, reviews, ranks):
    # Create matrix: [[], [], [], [], []]
    matrix = []
    for i in range(0, 5):
        matrix.append([])

    for i in range(len(reviews)):
        cont = 0
        polarity = 0
        for j in range(len(reviews[i])):
            if reviews[i][j] in dictionary:
                polarity = polarity +
                ↪ dictionary[reviews[i][j]]
                cont = cont + 1

        ans = 0
        if cont != 0:
            ans = polarity / cont

```

```

        rank = ranks[i] - 1
        matrix[rank].append(round(ans, 2))
    return matrix

# Getting polarity of every rank
def getTotalPolarity(m):
    result = [np.sum(row)/len(row) for row in m]
    return result

#####
#####
#####
# Get dictionary
fpath =
    ↪ '/Users/abiga/Desktop/AbiiSnn/GitHub/Natural-Language-Processing/Practice/25/ML-Sen
dictionary = getDictionary(fpath, 'utf-8') # Dictionary of tuples with
    ↪ polarity

# Get tokens by corpus
fpath =
    ↪ '/Users/abiga/Desktop/AbiiSnn/GitHub/Natural-Language-Processing/Practice/25/corpus
n = 4395
reviews = getText(fpath, 'ISO-8859-1', n) # List of tuples
cleanReviews = removeStopwords(reviews, 'english')
ranks = getRank(fpath, code, n)
matrix = getMatrix(dictionary, cleanReviews, ranks)
total = getTotalPolarity(matrix)

print("")
for i in range(len(matrix)):
    print("-----> RANK:", str(i+1), " # Reviews:", len(matrix[i]), "
        ↪ Total Polarity:", total[i])
    # print("\n")

print("")
for i in range(len(matrix)):
    print("-----> RANK:", str(i+1), " # Reviews:", len(matrix[i]), "
        ↪ Total Polarity:", total[i])
    print(matrix[i])
    print("\n")

```