## **Composición de la muestra balanceada para cada etiqueta**

Las etiquetas marcadas en rojo fueron excluidas por falta de datos suficientes para utilizar el algoritmo.

Español:

|  |  |  |  |
| --- | --- | --- | --- |
| **ETIQUETA** | **TOTAL**  **(2990 REGISTROS)** | | **MUESTRA BALANCEADA** |
| **NEGATIVO** | **POSITIVO** |
| V10\_PLACEMENT | 2735 | 255 | 510 |
| V11\_01\_Government\_Legislatures | 2101 | 889 | 1778 |
| V11\_02\_Campaigns\_Elections\_Politics | 2720 | 270 | 540 |
| V11\_03\_Police\_and\_Crime | 2694 | 296 | 592 |
| V11\_04\_Court | 2807 | 183 | 366 |
| V11\_05\_Defense\_Military\_National\_Security | 2933 | 57 | 114 |
| V11\_06\_Economy\_and\_Business | 2539 | 451 | 902 |
| V11\_07\_Education | 2908 | 82 | 164 |
| V11\_08\_Environment\_Sustainability\_Climate\_and\_Weather | 2901 | 89 | 178 |
| V11\_09\_Energy\_and\_Resource\_Development | 2966 | 24 | 48 |
| V11\_10\_Transportation | 2899 | 91 | 182 |
| V11\_11\_Housing\_Infrastructure\_and\_Public\_Works | 2959 | 31 | 62 |
| V11\_12\_Accidents\_and\_Natural\_Disasters | 2887 | 103 | 206 |
| V11\_13\_Health | 2141 | 849 | 1698 |
| V11\_14\_Religion\_and\_Churches | 2970 | 20 | 40 |
| V11\_15\_Labor\_and\_Employment | 2838 | 152 | 304 |
| V11\_16\_Demonstrations\_and\_Protests | 2889 | 101 | 202 |
| V11\_17\_Social\_Issues | 2855 | 135 | 270 |
| V11\_18\_Media\_and\_Communication | 2925 | 65 | 130 |
| V11\_19\_Sports | 2520 | 470 | 940 |
| V11\_20\_Science\_and\_Technology | 2838 | 152 | 304 |
| V11\_21\_Lifestyle | 2771 | 219 | 438 |
| V11\_22\_Culture | 2788 | 202 | 404 |
| V11\_23\_Entertainment\_and\_Celebrities | 2613 | 377 | 754 |
| V11\_24\_Others | 2753 | 237 | 474 |
| V14a\_STATEPOLITICALSOURCE | 1883 | 1107 | 2214 |
| V14b\_BUSINESSSOURCE | 2562 | 428 | 856 |
| V14c\_POLICESOURCE | 2868 | 122 | 244 |
| V14d\_LEGALSOURCE | 2816 | 174 | 348 |
| V14e\_MILITARYSOURCE | 2952 | 38 | 76 |
| V14f\_HEALTHSOURCE | 2647 | 343 | 686 |
| V14g\_EDUCATIONALSOURCE | 2760 | 230 | 460 |
| V14h\_CIVILSOCIETYSOURCE | 2785 | 205 | 410 |
| V14i\_RELIGIONSOURCE | 2971 | 19 | 38 |
| V14j\_CITIZENSOURCE | 2620 | 370 | 740 |
| V14k\_MEDIASOURCE | 2416 | 574 | 1148 |
| V14l\_SPORTSSOURCE | 2699 | 291 | 582 |
| V14m\_CELEBRITYSOURCE | 2661 | 329 | 658 |
| V14n\_ANONYMOUSSOURCE | 2916 | 74 | 148 |
| V14o\_OTHERSOURCE | 2848 | 142 | 284 |
| V17\_EXPERTSOURCE | 2667 | 323 | 646 |
| V18\_WITNESSSOURCE | 2714 | 276 | 552 |
| V19a\_OPINION | 2427 | 563 | 1126 |
| V19b\_INTERPRETATION | 2468 | 522 | 1044 |
| V19c\_CALLFOR | 2857 | 133 | 266 |
| V19d\_ADJECTIVES | 991 | 1999 | 1982 |
| V19e\_FIRSTPERSON | 2725 | 265 | 530 |
| V20a\_JUDICIALINF | 2726 | 264 | 528 |
| V20b\_DOUBTJOURN | 2885 | 105 | 210 |
| V20c\_DOUBTOTHER | 2730 | 260 | 520 |
| V20d\_CRITICJOURN | 2884 | 106 | 212 |
| V20e\_CRITICOTHER | 2377 | 613 | 1226 |
| V20f\_UNCOVERJOURN | 2912 | 78 | 156 |
| V20g\_UNCOVEROTHER | 2821 | 169 | 338 |
| V20h\_EXTERNALINV | 2836 | 154 | 308 |
| V20i\_INVREPORT | 2972 | 18 | 36 |
| V20j\_CONFLICT | 2988 | 2 | 4 |
| V21a\_DEFENSEACT | 2863 | 127 | 254 |
| V21b\_DEFPOLICIES | 2965 | 25 | 50 |
| V21c\_POSITIVEIMAGE | 2629 | 361 | 722 |
| V21d\_PROGRESS | 2979 | 11 | 22 |
| V21e\_COMPARISON | 2959 | 31 | 62 |
| V21f\_TRIUMPH | 2902 | 88 | 176 |
| V21g\_COUNTRYIMAGE | 2976 | 14 | 28 |
| V21h\_PATRIOTISM | 2968 | 22 | 44 |
| V22a\_IMPACTLIFE | 2576 | 414 | 828 |
| V22b\_TIPS1 | 2854 | 136 | 272 |
| V22c\_TIPS2 | 2822 | 168 | 336 |
| V22d\_CONSUMERINFO | 2590 | 400 | 800 |
| V22e\_CONSUMERADVICE | 2835 | 155 | 310 |
| V22f\_PERSASSISTANCE | 2951 | 39 | 78 |
| V23a\_PERSONALIZATION | 2198 | 792 | 1584 |
| V23b\_PRIVATELIFE | 2618 | 372 | 744 |
| V23c\_SENSATIONALISM | 2176 | 814 | 1628 |
| V23d\_EMOTIONS | 2295 | 695 | 1390 |
| V23e\_MORBIDITY | 2873 | 117 | 234 |
| V24a\_CITIZENREACTIONS | 2660 | 330 | 660 |
| V24b\_CITIZENDEMANDS | 2747 | 243 | 486 |
| V24c\_CREDCITIZENS | 2738 | 252 | 504 |
| V24d\_LOCALIMPACT | 2708 | 282 | 564 |
| V24e\_SOCIALIMPACT | 2789 | 201 | 402 |
| V24f\_EDUCATING | 2931 | 59 | 118 |
| V24g\_BACKGROUNDINFO | 2936 | 54 | 108 |
| V24h\_CITIZENQUESTIONS | 2965 | 25 | 50 |
| V24i\_INFOCITIZENSACT | 2904 | 86 | 172 |
| V24j\_SUPPORTMOVEMENTS | 2973 | 17 | 34 |

Inglés:

|  |  |  |  |
| --- | --- | --- | --- |
| **ETIQUETA** | **TOTAL**  **(1020 REGISTROS)** | | **MUESTRA BALANCEADA** |
| **NEGATIVO** | **POSITIVO** |
| V10\_PLACEMENT | 982 | 38 | 76 |
| V11\_01\_Government\_Legislatures | 715 | 305 | 610 |
| V11\_02\_Campaigns\_Elections\_Politics | 978 | 42 | 84 |
| V11\_03\_Police\_and\_Crime | 901 | 119 | 238 |
| V11\_04\_Court | 978 | 42 | 84 |
| V11\_05\_Defense\_Military\_National\_Security | 1007 | 13 | 26 |
| V11\_06\_Economy\_and\_Business | 895 | 125 | 250 |
| V11\_07\_Education | 992 | 28 | 56 |
| V11\_08\_Environment\_Sustainability\_Climate\_and\_Weather | 990 | 30 | 60 |
| V11\_09\_Energy\_and\_Resource\_Development | 1014 | 6 | 12 |
| V11\_10\_Transportation | 976 | 44 | 88 |
| V11\_11\_Housing\_Infrastructure\_and\_Public\_Works | 965 | 55 | 110 |
| V11\_12\_Accidents\_and\_Natural\_Disasters | 998 | 22 | 44 |
| V11\_13\_Health | 748 | 272 | 544 |
| V11\_14\_Religion\_and\_Churches | 1012 | 8 | 16 |
| V11\_15\_Labor\_and\_Employment | 997 | 23 | 46 |
| V11\_16\_Demonstrations\_and\_Protests | 1003 | 17 | 34 |
| V11\_17\_Social\_Issues | 918 | 102 | 204 |
| V11\_18\_Media\_and\_Communication | 971 | 49 | 98 |
| V11\_19\_Sports | 889 | 131 | 262 |
| V11\_20\_Science\_and\_Technology | 974 | 46 | 92 |
| V11\_21\_Lifestyle | 938 | 82 | 164 |
| V11\_22\_Culture | 964 | 56 | 112 |
| V11\_23\_Entertainment\_and\_Celebrities | 887 | 133 | 266 |
| V11\_24\_Others | 1010 | 10 | 20 |
| V14a\_STATEPOLITICALSOURCE | 591 | 429 | 858 |
| V14b\_BUSINESSSOURCE | 837 | 183 | 366 |
| V14c\_POLICESOURCE | 929 | 91 | 182 |
| V14d\_LEGALSOURCE | 967 | 53 | 106 |
| V14e\_MILITARYSOURCE | 1012 | 8 | 16 |
| V14f\_HEALTHSOURCE | 884 | 136 | 272 |
| V14g\_EDUCATIONALSOURCE | 870 | 150 | 300 |
| V14h\_CIVILSOCIETYSOURCE | 885 | 135 | 270 |
| V14i\_RELIGIONSOURCE | 1006 | 14 | 28 |
| V14j\_CITIZENSOURCE | 810 | 210 | 420 |
| V14k\_MEDIASOURCE | 767 | 253 | 506 |
| V14l\_SPORTSSOURCE | 909 | 111 | 222 |
| V14m\_CELEBRITYSOURCE | 921 | 99 | 198 |
| V14n\_ANONYMOUSSOURCE | 928 | 92 | 184 |
| V14o\_OTHERSOURCE | 986 | 34 | 68 |
| V17\_EXPERTSOURCE | 694 | 326 | 652 |
| V18\_WITNESSSOURCE | 758 | 262 | 524 |
| V19a\_OPINION | 924 | 96 | 192 |
| V19b\_INTERPRETATION | 640 | 380 | 760 |
| V19c\_CALLFOR | 990 | 30 | 60 |
| V19d\_ADJECTIVES | 689 | 331 | 662 |
| V19e\_FIRSTPERSON | 957 | 63 | 126 |
| V20a\_JUDICIALINF | 853 | 167 | 334 |
| V20b\_DOUBTJOURN | 980 | 40 | 80 |
| V20c\_DOUBTOTHER | 863 | 157 | 314 |
| V20d\_CRITICJOURN | 991 | 29 | 58 |
| V20e\_CRITICOTHER | 684 | 336 | 672 |
| V20f\_UNCOVERJOURN | 991 | 29 | 58 |
| V20g\_UNCOVEROTHER | 900 | 120 | 240 |
| V20h\_EXTERNALINV | 907 | 113 | 226 |
| V20i\_INVREPORT | 955 | 65 | 130 |
| V20j\_CONFLICT | 1020 | 0 | 0 |
| V21a\_DEFENSEACT | 998 | 22 | 44 |
| V21b\_DEFPOLICIES | 1015 | 5 | 10 |
| V21c\_POSITIVEIMAGE | 970 | 50 | 100 |
| V21d\_PROGRESS | 1015 | 5 | 10 |
| V21e\_COMPARISON | 1018 | 2 | 4 |
| V21f\_TRIUMPH | 1014 | 6 | 12 |
| V21g\_COUNTRYIMAGE | 1017 | 3 | 6 |
| V21h\_PATRIOTISM | 1016 | 4 | 8 |
| V22a\_IMPACTLIFE | 872 | 148 | 296 |
| V22b\_TIPS1 | 970 | 50 | 100 |
| V22c\_TIPS2 | 953 | 67 | 134 |
| V22d\_CONSUMERINFO | 933 | 87 | 174 |
| V22e\_CONSUMERADVICE | 991 | 29 | 58 |
| V22f\_PERSASSISTANCE | 990 | 30 | 60 |
| V23a\_PERSONALIZATION | 650 | 370 | 740 |
| V23b\_PRIVATELIFE | 885 | 135 | 270 |
| V23c\_SENSATIONALISM | 812 | 208 | 416 |
| V23d\_EMOTIONS | 724 | 296 | 592 |
| V23e\_MORBIDITY | 944 | 76 | 152 |
| V24a\_CITIZENREACTIONS | 854 | 166 | 332 |
| V24b\_CITIZENDEMANDS | 944 | 76 | 152 |
| V24c\_CREDCITIZENS | 949 | 71 | 142 |
| V24d\_LOCALIMPACT | 911 | 109 | 218 |
| V24e\_SOCIALIMPACT | 844 | 176 | 352 |
| V24f\_EDUCATING | 982 | 38 | 76 |
| V24g\_BACKGROUNDINFO | 858 | 162 | 324 |
| V24h\_CITIZENQUESTIONS | 1005 | 15 | 30 |
| V24i\_INFOCITIZENSACT | 983 | 37 | 74 |
| V24j\_SUPPORTMOVEMENTS | 1016 | 4 | 8 |

## Mejores y peores métricas en las etiquetas

El procedimiento fue comparar qué método de entrenamiento-evaluación (designado como [1] para el método 80-20 y [2] para el método de validación cruzada) tuvo mejores métricas en cada etiqueta, conformando para cada una un conjunto de mejores (5) y peores (5). De los 5 mejores se obtuvo el de la métrica más alta (para cada métrica) y de los 5 peores se obtuvo el de la métrica más baja (para cada métrica).

Finalmente, se destacaron, para cada designación Vxx (con xx ={11,14,19,20,21,22,23,24}), los 3 valores más altos y los 3 valores más bajos.