Sprawozdanie

Sztuczna Inteligencja i Inżynieria Wiedzy  
 Uczenie maszynowe

Imię i nazwisko autora: Adam Krzemiński

Nr indeksu: 246750

Data: 18.06.2021 r.

Semestr studiów: 6

Prowadzący laboratorium: mgr. inż. Jan Jakubik

Spis treści

[1. Opis analizowanego zbioru danych 3](#_Toc73100308)

[2. Opis i cel zaimplementowanego systemu maszynowego uczenia się 3](#_Toc73100309)

[3. Analiza eksploracyjna danych 3](#_Toc73100310)

[a. Liczba recenzji w zależności od klasy 3](#_Toc73100311)

[b. Liczba recenzji w zależności od oceny 4](#_Toc73100312)

[c. Średnia liczba słów recenzji w zależności od klasy 5](#_Toc73100313)

[d. Najczęściej występujące słowa w klasach 6](#_Toc73100314)

[4. Przeprowadzone badania 7](#_Toc73100315)

[5. Podsumowanie 7](#_Toc73100316)

# Opis analizowanego zbioru danych

# Opis i cel zaimplementowanego systemu maszynowego uczenia się

# Analiza eksploracyjna danych

## Liczba recenzji w zależności od klasy

|  |  |  |
| --- | --- | --- |
| Podział | Klasa | Liczba recenzji |
| 3 Klasy | 0 | 1197 |
| 1 | 1915 |
| 2 | 1894 |
| 4 Klasy | 0 | 615 |
| 1 | 1553 |
| 2 | 1998 |
| 3 | 840 |

## Liczba recenzji w zależności od oceny

## Średnia liczba słów recenzji w zależności od klasy

|  |  |  |
| --- | --- | --- |
| Podział | Klasa | Średnia liczba słów |
| 3 Klasy | 0 | 363 |
| 1 | 366 |
| 2 | 407 |
| 4 Klasy | 0 | 376 |
| 1 | 356 |
| 2 | 381 |
| 3 | 428 |

## Najczęściej występujące słowa w klasach

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Podział na 3 klasy | | | | | |
| 0 klasa | | 1 klasa | | 2 klasa | |
| **word** | **count** | **word** | **count** | **word** | **count** |
| film | 4080 | film | 5951 | film | 5897 |
| movie | 2381 | movie | 3293 | movie | 3488 |
| like | 1609 | like | 2340 | like | 2248 |
| story | 1156 | story | 1998 | story | 1994 |
| just | 1075 | time | 1470 | director | 1549 |
| time | 880 | just | 1431 | films | 1533 |
| director | 848 | director | 1337 | time | 1493 |
| good | 788 | good | 1277 | just | 1493 |
| films | 773 | little | 1252 | picture | 1293 |
| bad | 732 | characters | 1194 | characters | 1276 |
| make | 711 | films | 1144 | best | 1214 |
| characters | 682 | picture | 1033 | good | 1105 |
| little | 639 | make | 944 | little | 1090 |
| script | 616 | character | 904 | life | 1046 |
| look | 517 | way | 900 | character | 1023 |
| way | 510 | does | 855 | does | 924 |
| does | 509 | script | 840 | way | 924 |
| character | 505 | scenes | 839 | make | 883 |
| picture | 498 | look | 785 | performance | 834 |
| audience | 482 | best | 753 | script | 812 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Podział na 4 klasy | | | | | | | |
| 0 klasa | | 1 klasa | | 2 klasa | | 3 klasa | |
| **word** | **count** | **word** | **count** | **word** | **count** | **word** | **count** |
| film | 6173 | film | 11081 | film | 11773 | film | 22827 |
| movie | 3731 | movie | 5909 | movie | 7116 | movie | 1566 |
| like | 2465 | like | 4308 | like | 4644 | like | 980 |
| story | 1684 | story | 3686 | story | 3944 | story | 975 |
| just | 1653 | just | 2638 | director | 3093 | director | 765 |
| time | 1359 | time | 2595 | films | 3022 | films | 711 |
| director | 1301 | director | 2361 | time | 2884 | time | 708 |
| good | 1215 | good | 2236 | just | 2578 | best | 657 |
| bad | 1205 | little | 2229 | picture | 2574 | just | 653 |
| films | 1205 | characters | 2145 | characters | 2524 | picture | 608 |
| make | 1100 | films | 2044 | good | 2416 | character | 574 |
| characters | 1009 | picture | 1765 | little | 2342 | life | 554 |
| little | 947 | make | 1751 | best | 2177 | good | 471 |
| script | 945 | character | 1620 | life | 2054 | characters | 461 |
| look | 791 | way | 1576 | character | 1884 | way | 445 |
| way | 774 | does | 1556 | does | 1849 | little | 442 |
| does | 773 | script | 1537 | way | 1829 | does | 412 |
| character | 771 | scenes | 1490 | make | 1811 | make | 412 |
| picture | 749 | look | 1412 | performance | 1664 | year | 387 |
| audience | 736 | better | 1323 | script | 1654 | great | 386 |

# Przeprowadzone badania

# Podsumowanie