

Taller 3: Análisis Inteligente de Datos

Felipe Carmona

Sebastián Vergara

14 de Julio del 2016

1 Reducción de Dimensionalidad para Clasificación

La cantidad de datos por *dataframe* es el siguiente:

- Entrenamiento: 528 datos.
- Testing / Pruebas: 462 datos.

Los promedios obtenidos fueron los siguientes:

- Entrenamiento: 4,8 palabras para cada clase.
- Testing / Pruebas: 4,2 palabras para cada clase.

Luego se procede a normalizar los datos de ambos *dataframes*, para evitar problemas con las dimensiones de cada ítem.

A continuación se realiza una reducción de dimensionalidad de 10 a 2 dimensiones utilizando la técnica PCA. Luego se grafica ambas dimensiones y se diferencia cada clase con un color distinto, utilizando la paleta *Set1*.

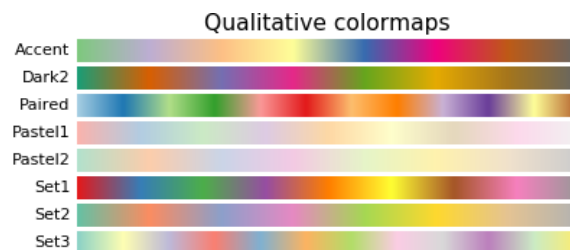


Figure 1: *Colormap* cualitativo



Figure 2: Disminución dimensión con PCA

Luego, se procede también a realizar una reducción de dimensionalidad pero esta vez utilizando la técnica de Fisher o disminución LDA.



Figure 3: Disminución dimensión con LDA

Se puede observar una gran diferencia en cuanto a la separabilidad de las clases en ambos gráficos. En el caso de PCA se observa que todos los puntos se encuentran dispersos en todo el espacio, impidiendo así generar grupos de puntos con colores similares (mismas clases). En cambio cuando se utiliza LDA, se evidencia que existe más separabilidad de los punto que son de la misma clase, logrando así poder generar agrupaciones de las distintas clases. El método más eficiente para realizar una reducción de dimensionalidad, para el caso estudiado, será LDA o Discriminante de Fisher.

Luego se pide comparar el desempeño de LDA, QDA y un modelo K-NN sin reducir dimensionalidad, como medida de comparación se utilizará el promedio de la precisión de cada modelo. A continuación se presentan los resultados:

| | Entranamiento | Testing |
|------|---------------|---------|
| LDA | 0.68 | 0.45 |
| QDA | 0.99 | 0.42 |
| K-NN | 0.93 | 0.49 |

Table 1: Resultados modelos sin reducir dimensionalidad

Luego se procede a realizar un análisis del modelo K-NN. En este caso los datos aún siguen con todas sus dimensiones. El valor de K varió desde 1 hasta el largo de los datos de entrenamiento que son 528.

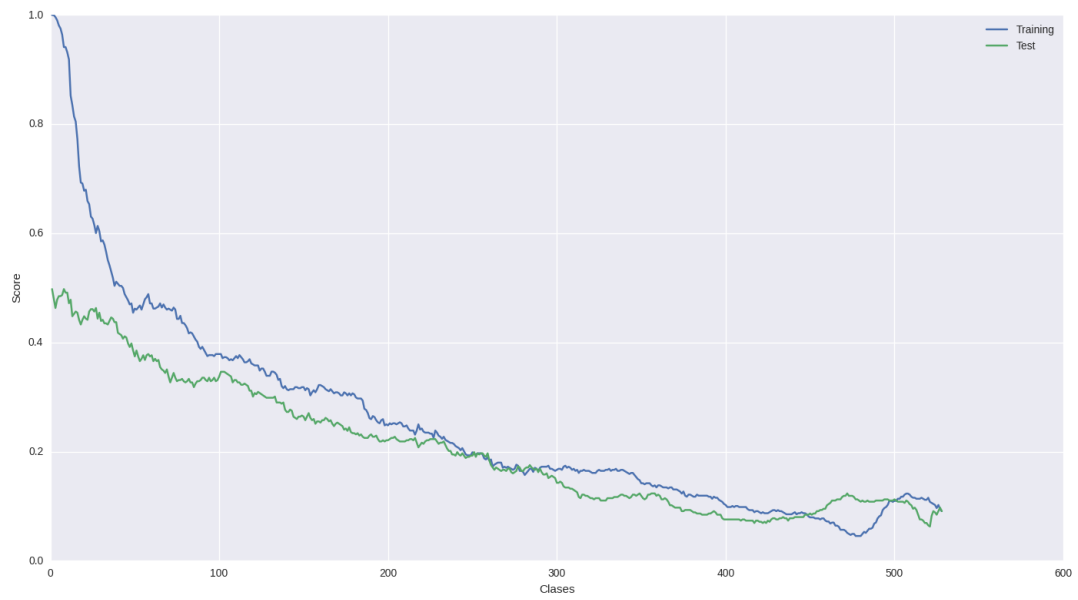


Figure 4: Precisión de KNN en función del parámetro K

Se observa que la precisión del modelo empeora a medida que el valor de K aumenta, esto puede deberse a que se presencia el fenómeno de *Overfitting*.

Para continuar la actividad se pide realizar nuevamente reducción de dimensionalidad para luego aplicar los modelos vistos anteriormente. La reducción en este caso se realizó con PCA. A continuación se presentan los resultados de la precisión de los modelos en función de la cantidad de dimensiones que tienen los datos.

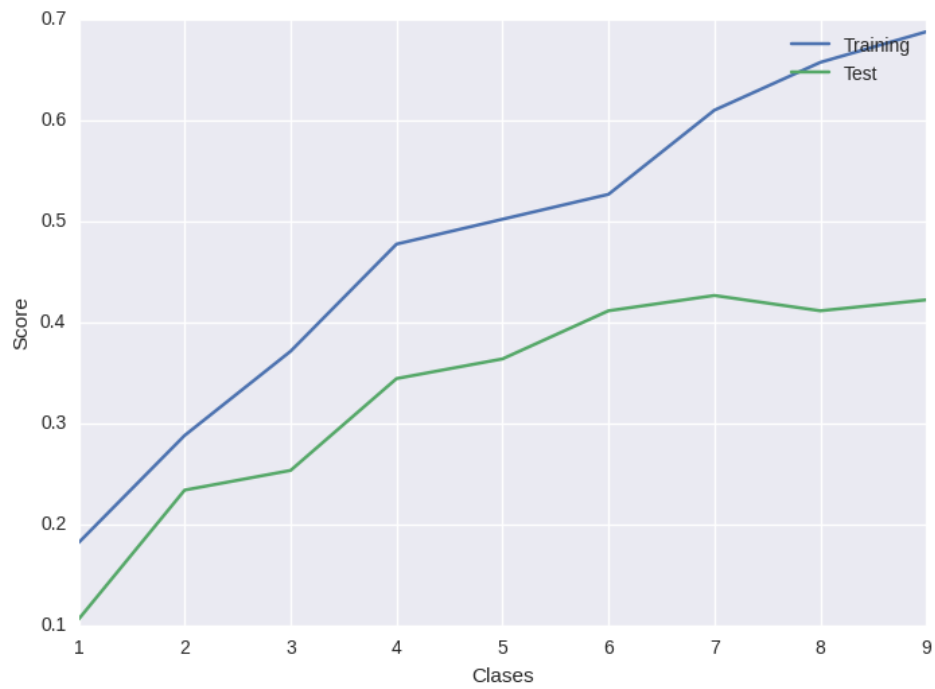


Figure 5: Precisión de LDA en función de la cantidad de dimensiones, caso PCA

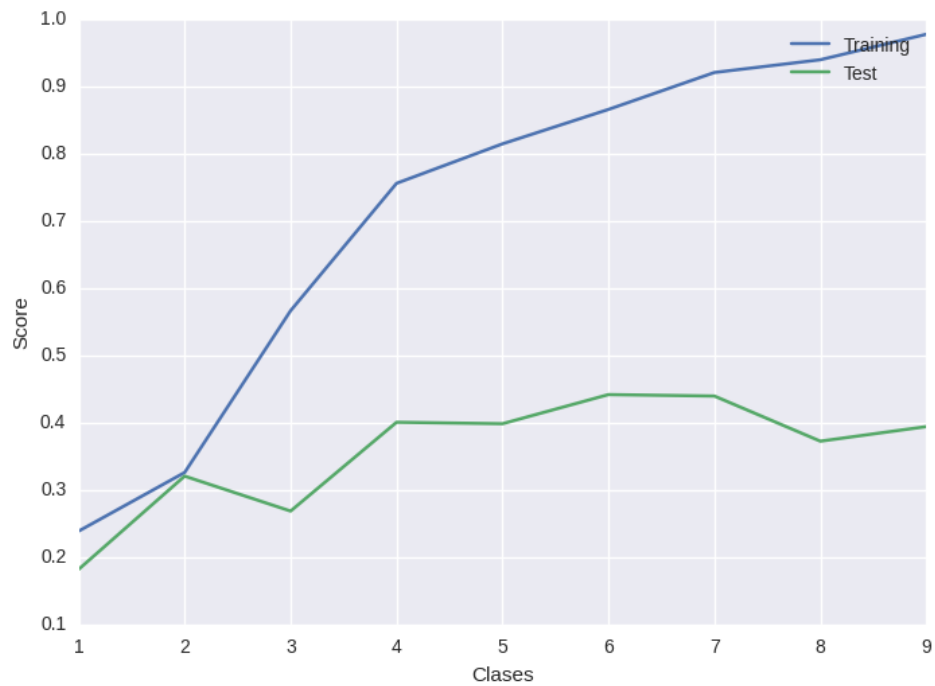


Figure 6: Precisión de QDA en función de la cantidad de dimensiones, caso PCA

Para K-NN se utilizó un K igual a 10.

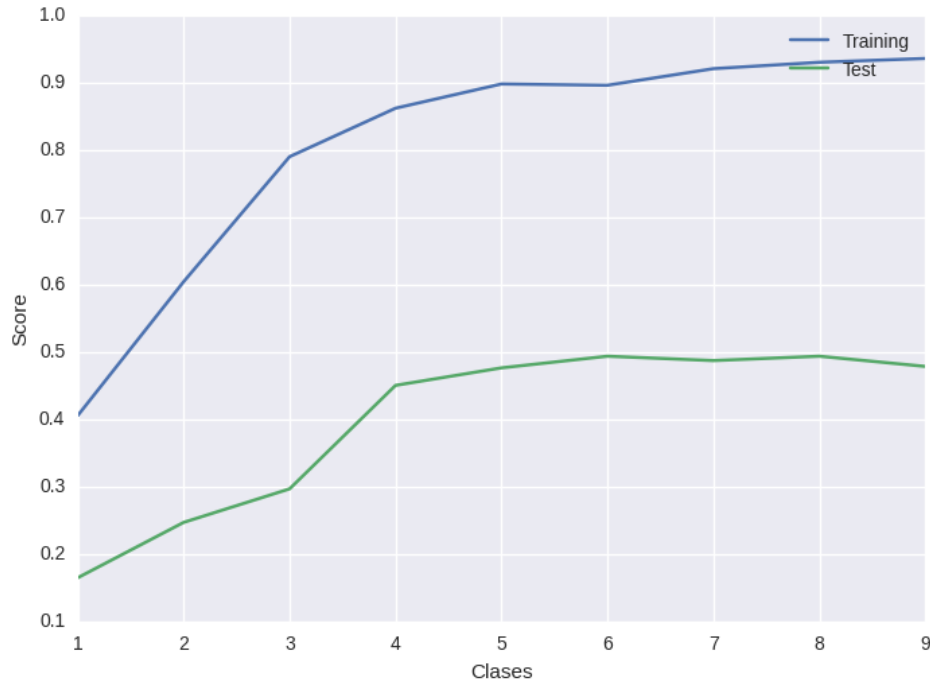


Figure 7: Precisión de K-NN en función de la cantidad de dimensiones, caso PCA

Se observa que en LDA, QDA y K-NN la precisión del modelo con los datos de entrenamiento aumenta monotonamente, en cambio para los datos de testing se tiene que existe una dimensión que maximiza la precisión del modelo y que se encuentra entre 1 y 9. Para LDA y QDA se tiene que el modelo presenta mejores resultados con 7 dimensiones. Para K-NN se observa que el número de dimensiones que otorga mayor precisión es 6. El orden de qué modelo da mejor precisión que los demás es: K-NN, QDA y LDA.

Para ambos los 3 modelos se observa que con los datos de testing se tienen buenas precisiones, tanto así que la reducción de dimensionalidad entrega mejor precisión que con todas las dimensiones.

A continuación se presentan los mismos gráficos que la parte anterior, con la diferencia que la reducción de dimensionalidad se realiza en base a la técnica LDA.

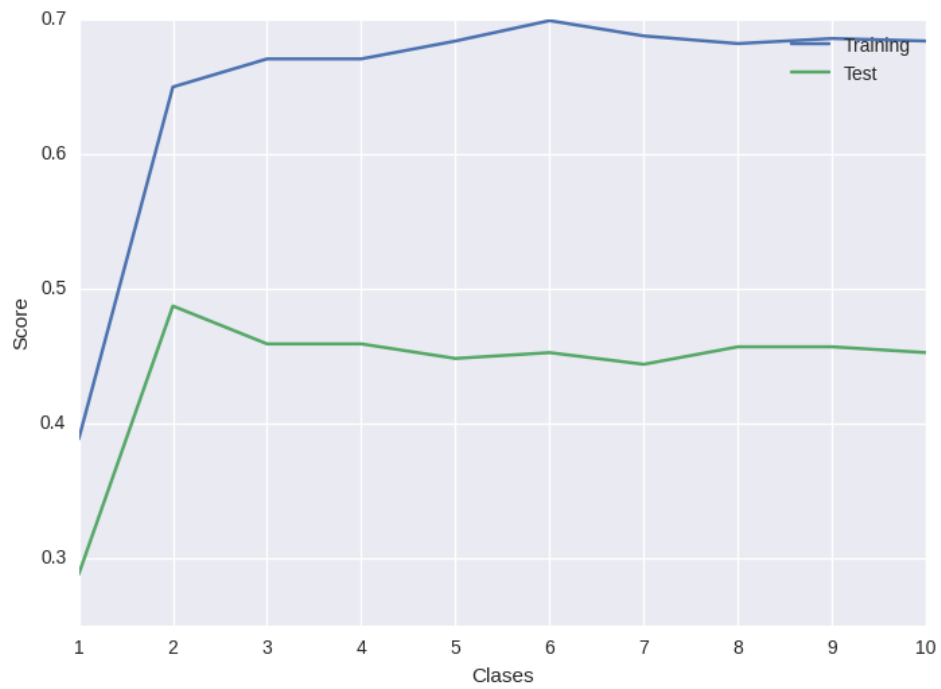


Figure 8: Precisión de LDA en función de la cantidad de dimensiones, caso LDA

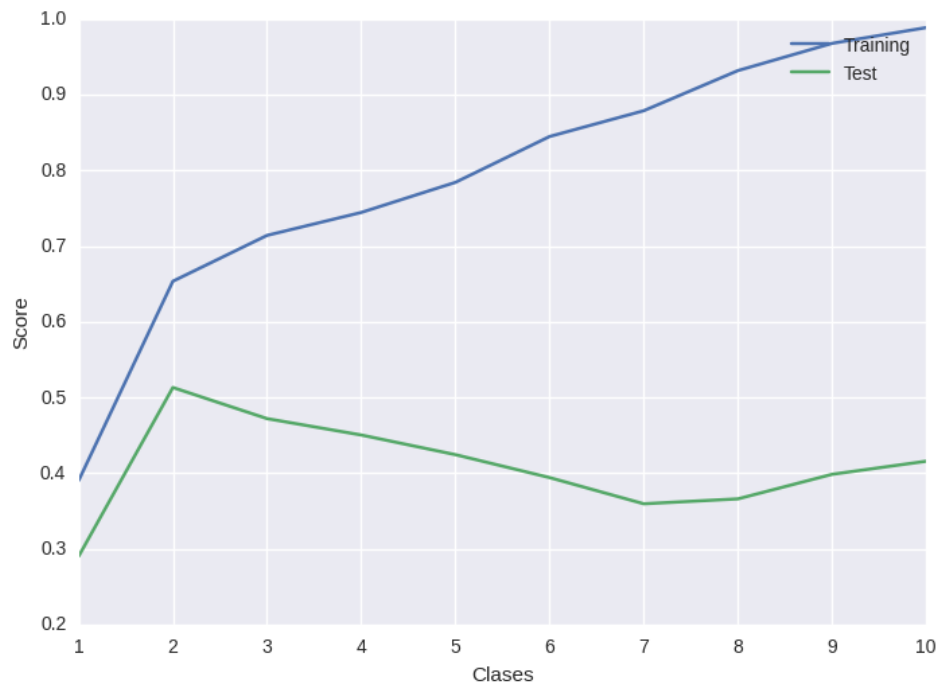


Figure 9: Precisión de QDA en función de la cantidad de dimensiones, caso LDA

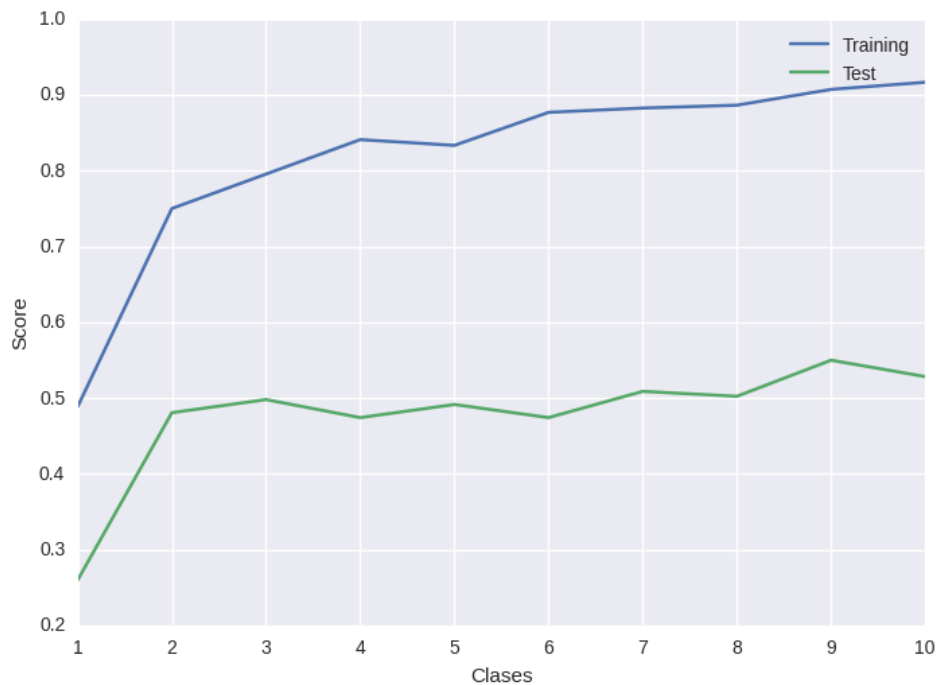


Figure 10: Precisión de K-NN en función de la cantidad de dimensiones, caso LDA

Al igual que en los gráficos anteriores, donde se realizó reducción de dimensionalidad con PCA, se puede observar que la precisión en los datos de entrenamiento aumentan a medida que se aumenta la dimensión (salvo para el caso de K-NN cuando aumentan las dimensiones de 4 a 5). Los mejores resultados de precisión se encuentran cuando se reduce hasta 2 dimensiones, para el caso de LDA y QDA. Para el caso de K-NN se encuentra que la mejor precisión está cuando se dejan 9 dimensiones.

Comparando con los resultados de disminuir dimensionalidad con PCA, se puede observar que LDA ofrece mejores resultados, tanto así, que la precisión es más de un 50%. En el caso de PCA las precisiones no superaba el 50%.

2 Análisis de Opiniones sobre Películas

2.1 Descripción de datos

Se dispone de dos dataframes, uno de entrenamiento y otro de prueba. Cada uno de ellos posee 3554 registros. A su vez, cada registro es descrito por dos características: *Sentimiento*, cuyo valor puede ser +1 (opinión positiva) ó -1 (opinión negativa) y *texto*, que contiene la opinión del espectador.

2.2 Preprocesamiento de texto: Stemming

A continuación se listan, los resultados obtenidos utilizando la función *word_extractor*. Primero se trabaja con stemming, haciendo uso del algoritmo de Porter:

I love to eat cake: love eat cake
I love eating cake: love eat cake
I loved eating the cake: love eat cake
I do not love eating cake: love eat cake
I don't love eating the cake: n't love eat cake
If it's so simple, why haven't you done it already?: 's simpl , whi n't done already

If you're good at something, never do it for free: 're good someth , never free
Well today I found out what Batman can't do: well today found batman ca n't

Se obtienen resultados pobres. Ello se manifiesta principalmente en que para las cuatro primeras expresiones, el resultado final es el mismo, aún cuando algunas de ellas tienen significados opuestos entre sí. Además, muchas de las expresiones generadas contienen tokens que no tiene significado alguno y se reducen tokens que no son verbos, como *something* y *already*

Sin el uso de stemming, los resultados son los siguientes:

I love to eat cake: love eat cake
I love eating cake: love eating cake
I loved eating the cake: love eating cake
I do not love eating cake: love eating cake
I don't love eating the cake: n't love eating cake
If it's so simple, why haven't you done it already?: 's simple , n't done already
If you're good at something, never do it for free: 're good something , never free
Well today I found out what Batman can't do: well today found batman ca n't

Se observa que los verbos no son reducidos a su tronco léxico base, pero ello favorece que no se reduzcan tokens que no corresponde reducir. Sin embargo, sigue siendo imposible distinguir entre una frase determinada y su negación. Se nota una mayor coherencia en las expresiones resultantes.

2.3 Preprocesamiento de texto: Lematización

Considerando las mismas expresiones de la sección anterior, se obtienen los siguientes resultados:

I love to eat cake: love eat cake
I love eating cake: love eating cake
I loved eating the cake: loved eating cake
I do not love eating cake: love eating cake
I don't love eating the cake: n't love eating cake
If it's so simple, why haven't you done it already?: 's simple , n't done already
If you're good at something, never do it for free: 're good something , never free
Well today I found out what Batman can't do: well today found batman ca n't

Los resultados son prácticamente idénticos a los obtenidos sin realizar stemming, excepto por la tercera expresión, donde la forma verbal *loved* se mantiene, pero no marca una diferencia significativa.

En resumen, se puede ver que, aunque tanto stemming como lematización tienen sus ventajas y desventajas, lematización obtiene mejores resultados, o en su defecto, la no utilización de stemming.

2.4 Construcción de vocabulario

Considerando tanto los datos de entrenamiento como los de prueba, es posible formar un vocabulario compuesto de 9811 tokens. A partir de este, se crea un *top ten*, de acuerdo a la frecuencia de cada uno (tabla 2).

Table 2: Top ten de tokens, según su frecuencia.

| Token | Frecuencia |
|------------------|------------|
| <i>film</i> | 581 |
| <i>movie</i> | 567 |
| <i>one</i> | 259 |
| <i>ha</i> | 246 |
| <i>like</i> | 239 |
| <i>story</i> | 204 |
| <i>character</i> | 178 |
| <i>time</i> | 176 |
| <i>make</i> | 167 |

2.5 Desempeño de clasificadores

A continuación, se muestra el desempeño de diversos modelos de clasificación aplicados sobre los datos provistos. Para esto, se expondrá un reporte por cada modelo. Dicho reporte indica la precisión, el recall, el valor-F (también conocido como F1-score) y el soporte de cada clase.

2.5.1 Clasificador Bayesiano Ingenuo Binario

A continuación, se muestran los resultados obtenidos para cada caso:

2.5.1.1 Caso 1: Filtrando stopwords y usando lematización

Precisión sobre datos de entrenamiento: 0,952

Precisión sobre datos de prueba: 0,751

Table 3: Caso 1: Resultados filtrando stopwords y usando lematización

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|-------------|-------------|-------------|-------------|
| +1 | 0,73 | 0,80 | 0,76 | 1784 |
| -1 | 0,78 | 0,71 | 0,74 | 1770 |
| Promedio / Total | 0,75 | 0,75 | 0,75 | 3554 |

2.5.1.2 Caso 2: Sin filtrar stopwords y usando lematización

Precisión sobre datos de entrenamiento: 0,949

Precisión sobre datos de prueba: 0,757

Table 4: Caso 2: Sin filtrar stopwords y usando lematización

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|-------------|-------------|-------------|-------------|
| +1 | 0,74 | 0,80 | 0,77 | 1784 |
| -1 | 0,78 | 0,72 | 0,75 | 1770 |
| Promedio / Total | 0,76 | 0,76 | 0,76 | 3554 |

Luego, al usar lematización, se obtiene una mayor precisión sobre el set de entrenamiento si se filtran stopwords, pero la precisión es mayor sobre el set de prueba en el caso en que no se filtran. De todas maneras, las diferencias son mínimas.

2.5.1.3 Caso 3: Filtrando stopwords y usando stemming

Precisión sobre datos de entrenamiento: 0,930

Precisión sobre datos de prueba: 0,758

Table 5: Caso 3: Resultados filtrando stopwords y usando stemming

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|------------------|---------------|----------------|----------------|
| +1 | 0,74 | 0,79 | 0,77 | 1784 |
| -1 | 0,77 | 0,73 | 0,75 | 1770 |
| Promedio / Total | 0,76 | 0,76 | 0,76 | 3554 |

Así, se obtiene una mayor precisión sobre el set de entrenamiento al usar lematización, pero la precisión sobre el set de prueba es mayor con stemming.

Si se toman en cuenta la precisión y recall por cada clase (en cada caso), puede decirse que los mejores resultados se consiguen al usar lematización y no filtrar stopwords.

2.5.1.4 Análisis de predicciones

Se han tomado cinco textos y se muestra la predicción sobre cada uno (se considera caso 1):

Texto: 'a' for creativity but comes across more as a sketch for a full-length comedy .

Predicción -1: 0,96

Predicción +1: 0,04

La opinión es mixta, pero el clasificador la considera más cercana a una opinión negativa.

Texto: every once in a while , a movie will come along that turns me into that annoying specimen of humanity that i usually dread encountering the most - the fanboy

Predicción -1: 0,94

Predicción +1: 0,06

La opinión es algo ambigua. Sin embargo, el clasificador la considera más bien una opinión negativa.

Texto: it just goes to show , an intelligent person isn't necessarily an admirable storyteller .

Predicción -1: 0,53

Predicción +1: 0,47

Se ve que la opinión es negativa, pero el clasificador la considera más cercana a una opinión mixta.

Texto: the movie is for fans who can't stop loving anime , and the fanatical excess built into it .

Predicción -1: 0,96

Predicción +1: 0,04

Es una opinión más bien neutra, pero el clasificador la asocia más como una crítica negativa.

Texto: there is truth here

Predicción -1: 0,44

Predicción +1: 0,56

Es una opinión ambigua, por lo que resulta apropiado que la predicción del clasificador sea mixta.

2.5.2 Clasificador Bayesiano Ingenuo Multinomial

A continuación, se muestran los resultados obtenidos para cada caso:

2.5.2.1 Caso 1: Filtrando stopwords y usando lematización

Precisión sobre datos de entrenamiento: 0,955

Precisión sobre datos de prueba: 0,755

Table 6: Caso 1: Resultados filtrando stopwords y usando lematización

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|------------------|---------------|----------------|----------------|
| +1 | 0,75 | 0,77 | 0,76 | 1784 |
| -1 | 0,76 | 0,74 | 0,75 | 1770 |
| Promedio / Total | 0,76 | 0,76 | 0,76 | 3554 |

2.5.2.2 Caso 2: Sin filtrar stopwords y usando lematización

Precisión sobre datos de entrenamiento: 0,949

Precisión sobre datos de prueba: 0,757

Table 7: Caso 2: Sin filtrar stopwords y usando stemming

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|------------------|---------------|----------------|----------------|
| +1 | 0,75 | 0,78 | 0,76 | 1784 |
| -1 | 0,77 | 0,74 | 0,75 | 1770 |
| Promedio / Total | 0,76 | 0,76 | 0,76 | 3554 |

Luego, al usar lematización, se obtiene una mayor precisión sobre el set de entrenamiento si se filtran stopwords, pero la precisión es mayor sobre el set de prueba en el caso en que no se filtran. De todas maneras, las diferencias son mínimas.

2.5.2.3 Caso 3: Filtrando stopwords y usando stemming

Precisión sobre datos de entrenamiento: 0,933

Precisión sobre datos de prueba: 0,758

Table 8: Caso 3: Resultados filtrando stopwords y usando stemming

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|------------------|---------------|----------------|----------------|
| +1 | 0,74 | 0,79 | 0,77 | 1784 |
| -1 | 0,77 | 0,73 | 0,75 | 1770 |
| Promedio / Total | 0,76 | 0,76 | 0,76 | 3554 |

Así, se obtiene una mayor precisión sobre el set de entrenamiento al usar lematización, pero la precisión sobre el set de prueba es mayor con stemming.

Si se toman en cuenta la precisión y recall por cada clase (en cada caso), puede decirse que los mejores resultados se consiguen al usar lematización y no filtrar stopwords.

2.5.2.4 Análisis de predicciones

Se han tomado cinco textos y se muestra la predicción sobre cada uno (se considera caso 1):

Texto: passionate , irrational , long-suffering but cruel as a tarantula , helga figures prominently in this movie , and helps keep the proceedings as funny for grown-ups as for rugrats

Predicción -1: 0,04

Predicción +1: 0,96

La opinión es positiva, y así como le predice el clasificador.

Texto: an eccentric little comic/thriller deeply in love with its own quirky personality .

Predicción -1: 0,23

Predicción +1: 0,77

La opinión puede considerarse más bien negativa, pero el clasificador la considera más cercana a una opinión positiva.

Texto: rice never clearly defines his characters or gives us a reason to care about them .

Predicción -1: 0,94

Predicción +1: 0,06

La opinión es negativa, y así es como el clasificador lo predice.

Texto: after a while , the only way for a reasonably intelligent person to get through the country bears is to ponder how a whole segment of pop-music history has been allowed to get wet , fuzzy and sticky .

Predicción -1: 0,98

Predicción +1: 0,02

La opinión es negativa, y así es como lo predice el clasificador.

Texto: The music and the stars aren't enough to save the movie .

Predicción -1: 0,97

Predicción +1: 0,03

La opinión es negativa, y así es como el clasificador lo entiende.

2.5.3 Modelo de regresión logística regularizado

El efecto esperado del parámetro C es regularizar el modelo, es decir, encontrar valores para los coeficientes asociados a cada variable presente en la regresión, de tal manera que se minimice el error de predicción.

Analizando los resultados obtenidos para cada valor de C, se tiene:

Table 9: Rendimiento de modelo de regresión logística regularizado para distintos valores de C

| C | Precisión set entrenamiento | Precisión set de prueba |
|------|-----------------------------|-------------------------|
| 0,01 | 0,792 | 0,680 |
| 0,1 | 0,897 | 0,716 |
| 10 | 1 | 0,706 |
| 100 | 1 | 0,698 |

A partir de las métricas expuestas en la tabla 9, se elige el parámetro $C = 10$ como el más apropiado para realizar la regresión, dado que presenta un buen equilibrio entre precisión sobre el set de entrenamiento y precisión sobre el set de prueba. Luego, se utiliza este valor para examinar diferencias en el rendimiento en base a la consideración de stopwords y el uso de stemming o lematización. Se obtienen las siguientes métricas:

Table 10: Caso 1: Filtrando stopwords y usando lematización. $C = 10$

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|-------------|-------------|-------------|-------------|
| +1 | 0,70 | 0,72 | 0,71 | 1784 |
| -1 | 0,71 | 0,69 | 0,70 | 1770 |
| Promedio / Total | 0,71 | 0,71 | 0,71 | 3554 |

Precisión sobre datos de entrenamiento: 1

Precisión sobre datos de prueba: 0,706

Table 11: Caso 2: Sin filtrar stopwords y usando lematización. C = 10

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|------------------|---------------|----------------|----------------|
| +1 | 0,70 | 0,72 | 0,71 | 1784 |
| -1 | 0,71 | 0,69 | 0,70 | 1770 |
| Promedio / Total | 0,71 | 0,71 | 0,71 | 3554 |

Precisión sobre datos de entrenamiento: 1
 Precisión sobre datos de prueba: 0,728

Luego, la precisión sobre el set de entrenamiento es la misma, independiente de que se filtren o no stopwords. Sin embargo, la precisión sobre el set de prueba es mayor en el segundo caso.

Table 12: Caso 3: Filtrando stopwords y usando stemming. C = 10

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|------------------|---------------|----------------|----------------|
| +1 | 0,70 | 0,72 | 0,71 | 1784 |
| -1 | 0,71 | 0,69 | 0,70 | 1770 |
| Promedio / Total | 0,71 | 0,71 | 0,71 | 3554 |

Precisión sobre datos de entrenamiento: 0,999
 Precisión sobre datos de prueba: 0,725

Se observa que la precisión sobre ambos sets de datos es siempre superior al usar lematización, respecto a usar stemming.

Finalmente, se toman cinco textos aleatoriamente y se muestra la predicción sobre cada uno (caso 1):

Texto: entertains by providing good , lively company .
Predicción -1: 0,01
Predicción +1: 0,99

La opinión es positiva, y así es como lo entiende el clasificador.

Texto: k 19 stays afloat as decent drama/action flick
Predicción -1: 0,66
Predicción +1: 0,34

La opinión es más bien mixta, y así lo determina el clasificador.

Texto: in the end , white oleander isn't an adaptation of a novel . it's a flashy , star-splashed reduction .
Predicción -1: 0,12
Predicción +1: 0,88

La opinión es negativa, pero el clasificador la considera más cercana a una opinión positiva.

Texto: if we're to slap protagonist genevieve leplouff because she's french , do we have that same option to slap her creators because they're clueless and inept ?
Predicción -1: 0,52
Predicción +1: 0,48

Es una opinión que no expone claramente su polaridad, por lo que es adecuado que la clasificación sea mixta.

Texto: a sentimental mess that never rings true
Predicción -1: 0,93
Predicción +1: 0,07
 La opinión es negativa, y así lo entiende el clasificador.

2.5.4 SVM Lineal

El efecto esperado del parámetro C es regularizar el modelo, es decir, evitar que este se sobreajuste.

Analizando los resultados obtenidos para cada valor de C, se tiene:

Table 13: Rendimiento de modelo SVM regularizado para distintos valores de C

| C | Precisión set entrenamiento | Precisión set de prueba |
|------|-----------------------------|-------------------------|
| 0,01 | 0,889 | 0,717 |
| 0,1 | 0,989 | 0,714 |
| 10 | 1 | 0,687 |
| 100 | 1 | 0,682 |
| 1000 | 1 | 0,683 |

A partir de las métricas expuestas en la tabla 13, se elige el parámetro $C = 0,1$ como el más apropiado, dado que presenta un buen equilibrio entre precisión sobre el set de entrenamiento y precisión sobre el set de prueba. Luego, se utiliza este valor para examinar diferencias en el rendimiento en base a la consideración de stopwords y el uso de stemming o lematización. Se obtienen las siguientes métricas:

Table 14: Caso 1: Filtrando stopwords y usando lematización. $C = 0,1$

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|-------------|-------------|-------------|-------------|
| +1 | 0,70 | 0,75 | 0,72 | 1784 |
| -1 | 0,73 | 0,68 | 0,70 | 1770 |
| Promedio / Total | 0,71 | 0,71 | 0,71 | 3554 |

Precisión sobre datos de entrenamiento: 0,935
 Precisión sobre datos de prueba: 0,714

Table 15: Caso 2: Sin filtrar stopwords y usando lematización. $C = 0,1$

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|-------------|-------------|-------------|-------------|
| +1 | 0,71 | 0,74 | 0,73 | 1784 |
| -1 | 0,73 | 0,69 | 0,71 | 1770 |
| Promedio / Total | 0,71 | 0,71 | 0,71 | 3554 |

Precisión sobre datos de entrenamiento: 0,934
 Precisión sobre datos de prueba: 0,718

Luego, la precisión sobre el set de entrenamiento es prácticamente la misma, tanto para cuando se filtran como cuando no se filtran stopwords. Sin embargo, la precisión sobre el set de prueba es mayor cuando no se filtran stopwords.

Table 16: Caso 3: Filtrando stopwords y usando stemming. $C = 0,1$

| | Precisión | Recall | Valor-F | Soporte |
|-------------------------|-------------|-------------|-------------|-------------|
| +1 | 0,72 | 0,76 | 0,74 | 1784 |
| -1 | 0,74 | 0,70 | 0,72 | 1770 |
| Promedio / Total | 0,73 | 0,73 | 0,73 | 3554 |

Precisión sobre datos de entrenamiento: 0,922
Precisión sobre datos de prueba: 0,729

A diferencia de los modelos usados anteriormente, en este caso se obtiene una mayor precisión sobre los datos de prueba al usar stemming respecto a lematización.

Finalmente, se toman cinco textos aleatoriamente y se muestra la predicción sobre cada uno (caso 1):

Texto: what a bewilderingly brilliant and entertaining movie this is .
Predicción -1: 0,14
Predicción +1: 0,86

La opinión es positiva y el clasificador efectivamente lo considera así.

Texto: whether kiss is a future cult classic or destined to be completely forgotten is open to question , but the risk-takers in the crowd should check it out and form their own opinion .
Predicción -1: 0,37
Predicción +1: 0,63

Se puede considerar como un opinión neutra, y en parte así lo considera el clasificador, aunque más cercana a una opinión positiva.

Texto: life on the rez is no picnic : this picture shows you why .
Predicción -1: 0,47
Predicción +1: 0,53

Se puede considerar como una opinión neutral. El clasificador también lo considera así.

Texto: this movie is maddening . it conveys a simple message in a visual style that is willfully overwrought .
Predicción -1: 0,47
Predicción +1: 0,53

La opinión es más bien neutra, y así le predice el clasificador.

Texto: with wit and empathy to spare , waydowntown acknowledges the silent screams of workaday inertia but stops short of indulging its characters' striving solipsism .
Predicción -1: 0,60
Predicción +1: 0,40

Es una opinión mixta, y el clasificador así lo entiende.

2.6 Comparación de métodos de clasificación

Se evalúa cada método en base a la cantidad de clasificaciones True Positive (TP), False Positive (FP), True Negative (TN) y False Negative (FN), donde 'el positivo' es la clase +1, y 'el negativo' es la clase -1.

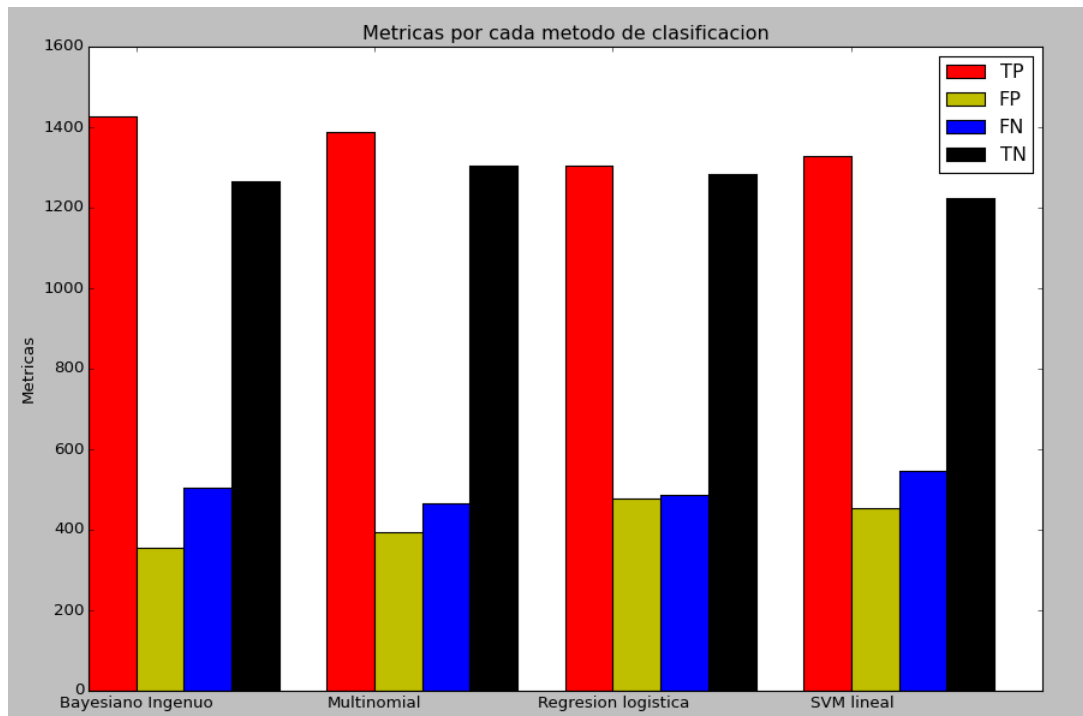


Figure 11: Comparación de métodos de clasificación.

Se observa que mejor predicción de clase +1 (TP) se logra con un modelo Bayesiano Ingenuo, mientras que la clase -1 (TN) obtiene una mejor predicción a partir de un modelo multinomial.