**1. El Modelo "Simple" fue más Inteligente que el "Complejo"**

Uno de los primeros hallazgos más contraintuitivos fue la comparación de modelos. Entrenamos un modelo de **Regresión Logística** (relativamente simple) y uno de **LightGBM** (un algoritmo avanzado y potente). La expectativa era que el modelo complejo superara al simple.

* **El Hallazgo Curioso:** La Regresión Logística ganó. Obtuvo una precisión mayor (81% vs 79%) y, lo que es más importante, fue significativamente mejor para identificar reseñas negativas.
* **¿Por qué es curioso?** Es un recordatorio clave en ciencia de datos: más complejidad no siempre es mejor. Para datos de texto como estos, un modelo lineal puede capturar las relaciones importantes de manera más efectiva y robusta que un modelo complejo que podría estar buscando patrones que no existen o que son puro ruido.

**2. El "Efecto Selma": Descubrimos un Fósil Lingüístico Cultural**

Cuando analizamos cómo cambiaban las palabras importantes a lo largo del tiempo, encontramos una anomalía fascinante en el período 2010-2015.

* **El Hallazgo Curioso:** La palabra **"selma"** apareció como uno de los predictores más fuertes de una reseña positiva. Obviamente, "selma" no es un adjetivo.
* **¿Por qué es curioso?** El modelo identificó un momento cultural. La película "Selma" (2014) fue tan aclamada por la crítica que se convirtió en un **punto de referencia**. Los críticos comenzaron a usarla como un estándar de calidad en otras reseñas (ej. "tan impactante como Selma", "no logra la resonancia de Selma"). El modelo fue lo suficientemente inteligente como para aprender que la mera mención de esta película en una reseña era un fuerte indicador de una discusión de cine de alta calidad, y por lo tanto, de un sentimiento positivo.

**3. El Lenguaje del Elogio Evoluciona, el del Fracaso es Eterno**

Al comparar las palabras clave a lo largo de 20 años, notamos una asimetría muy clara.

* **El Hallazgo Curioso:** El vocabulario para describir una mala película (dull, boring, fails, tedious) es increíblemente estable. Las palabras negativas de 2002 son casi las mismas que las de 2020. Sin embargo, el vocabulario para una buena película es más dinámico, incorporando términos nuevos y más específicos con el tiempo ('exquisite', 'brilliant', 'delightful').
* **¿Por qué es curioso?** Sugiere que las razones por las que fracasa una película son fundamentales y atemporales (es aburrida, la trama no funciona). En cambio, las formas de alabar una película se adaptan y evolucionan, quizás reflejando modas en el lenguaje o nuevos tipos de virtudes cinematográficas que se empiezan a valorar.

**4. Una Comedia Romántica puede ser tan Divisiva como el Cine de Autor**

Cuando buscamos las películas más polarizantes, la lista no solo incluyó películas de autor o de terror controvertidas, como uno esperaría.

* **El Hallazgo Curioso:** La comedia romántica **"Kate & Leopold"** (2001) y la película de autor **"Manderlay"** (2005) generaron estadísticamente el mismo nivel de división entre los críticos (un 50/50 perfecto entre 'Fresh' y 'Rotten').
* **¿Por qué es curioso?** Demuestra que la polarización no solo proviene de temas controversiales, sino también de la ejecución y el género. Una comedia que algunos críticos consideran encantadora y otros insoportablemente cursi puede generar tanto desacuerdo como una película experimental diseñada para provocar.

Estos hallazgos son el verdadero tesoro de tu proyecto. No solo demuestran que tu modelo funciona, sino que también revelan historias inesperadas sobre la cultura, el lenguaje y el cine.