Baahubali 2: Dissecting 75,000 Tweets to Uncover Audience Sentiments

5th May, 2017

BY BYOM KESH

Why did Katappa kill Baahubali?

Two years ago, not many would have foreseen this sentence capturing the imagination of the country like it has. Demolishing all regional barriers, the movie has grossed over INR 500 crores across the world in only its first three days.

While the first movie received lavish praise for its ambition, technical values, and story, the sequel, bogged by bloated expectations, has polarized the critics fraternity. Some critics compare the movie's computer graphics favorably to Hollywood productions like Lord of the Rings. Others find the movie lacking in pacing and plot.
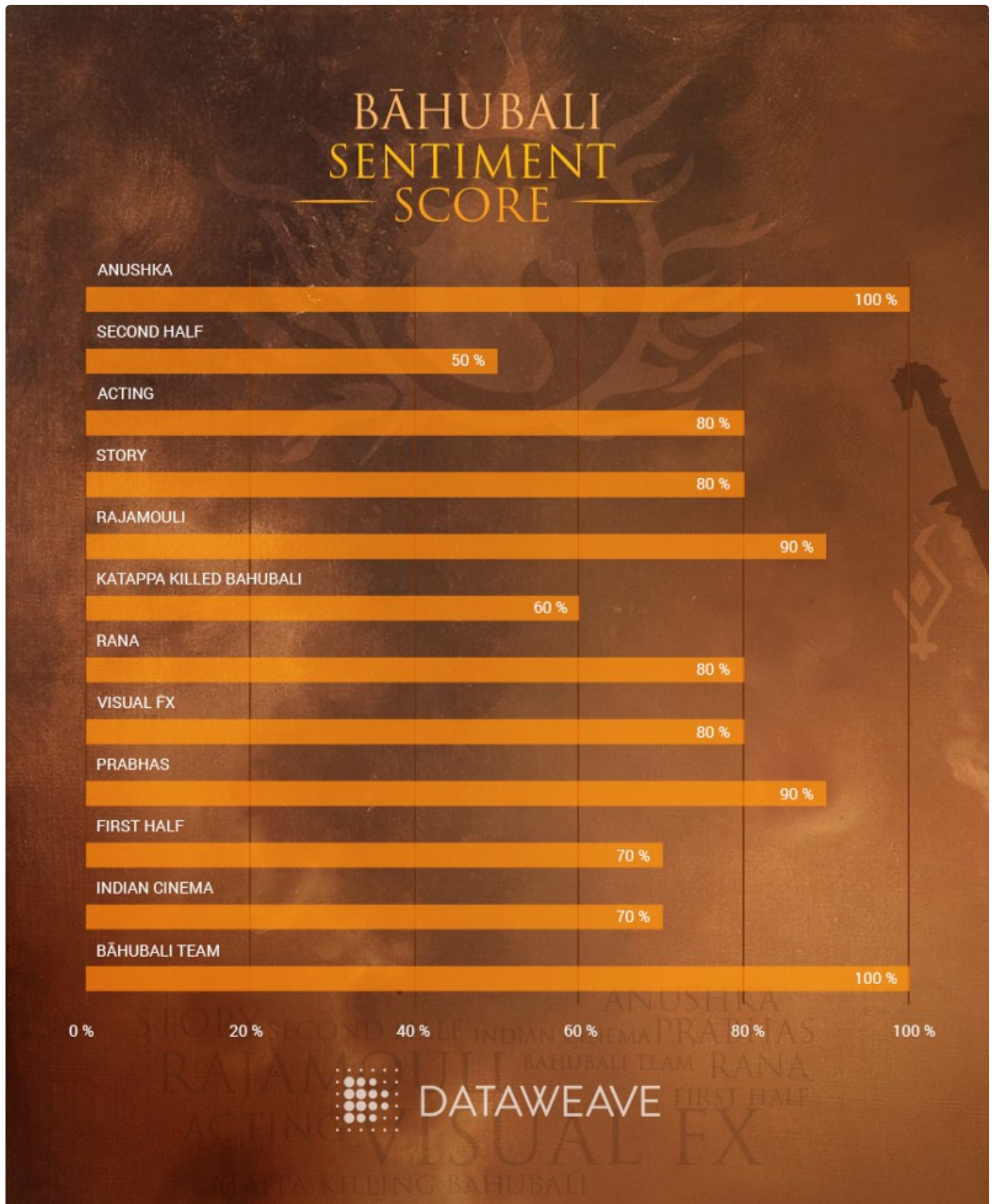
The masses, however, have reportedly lapped the movie up. Social media channels are brimming with opinions, and if one is to attempt finding out the aggregate views of audiences, Twitter is a good place to start.

At DataWeave, we ran our proprietary, AI-powered 'Sentiment Analysis' algorithm over all tweets about Baahubali 2 the first three days of its release, and observed some interesting insights.

Twitterati Reactions to Baahubali 2

Overall, the Twitterati's views on the movie were overwhelmingly positive. We analysed over **75,000 tweets** and identified the sentiments expressed on

several facets of the movie, such as, Visuals, Acting, Prabhas, etc. The following graphic indicates how the movie fared in some of these categories.

BĀHUBALI SENTIMENT SCORE

Category	Score
ANUSHKA	100 %
SECOND HALF	50 %
ACTING	80 %
STORY	80 %
RAJAMOULI	90 %
KATAPPA KILLED BAHUBALI	60 %
RANA	80 %
VISUAL FX	80 %
PRABHAS	90 %
FIRST HALF	70 %
INDIAN CINEMA	70 %
BĀHUBALI TEAM	100 %

0 %   20 %   40 %   60 %   80 %   100 %

DATAWEAVE

The Baahubali team, Anushka (actor), Rajamouli (director), and Prabhas (actor), are all perceived as huge positive influences on the movie. Rajamouli, specifically, met with almost universal approval for his dedication and execution. Several viewers cheered the movie on as a triumph of Indian cinema, one which has redefined the cinema landscape of the country. There was considerable praise for the story, Rana (actor), and acting performances, as well.

The not-so-positive sentiments were reserved for the reason behind Katappa killing Baahubali (no spoilers!), the visuals, and the second half of the movie. Many viewers found the second half to be slow, with unrealistic visuals and action sequences. For example, one of the tweets read:

"First half was good, but the second half is beyond Rajnikanth movies: humans uprooting trees!"

While these insights seem simple enough to understand, the technology to filter inevitably chaotic online content and extract meaningful information is incredibly complex.

Unearthing Meaning from Chaos

At DataWeave, we provide enterprises with Competitive Intelligence as a Service by aggregating and analyzing millions of unstructured data points on the web, across multiple sources. This enables businesses to better understand their competitive environment and make data-driven decisions to grow their business.

One of our solutions — Sentiment Analysis — helps brands study customer preferences at a product attribute level by analyzing customer reviews. We used the same technology to analyze the reaction of audiences globally to Baahubali 2. After data acquisition, this process consists of three steps –

Step 1: Features Extraction

To identify the "features" that reviewers are talking about, we first understand the syntactical structure of the tweets and separate words into nouns, verbs, adjectives, etc. This needs to account for complexities like synonyms, spelling errors, paraphrases, noise, etc. Our AI-based technology platform then uses various advanced techniques to generate a list of "uni-features" and "compound features" (more than one word for a feature).

Step 2: Identifying Feature-Opinion Pairs

Next, we identify the relationship between the feature and the opinion. One of the reasons this is challenging with twitter is, most of the time, twitter users treat grammar with utter disdain. Case in point:

"I saw the movie visuals awesome bad climax felt director unnecessarily dragged the second half"

In this case, the feature-opinion pairs are visuals: awesome, climax: bad, second half: unnecessarily dragged. Clearly, something as simple as attributing the nearest opinion-word to the feature is not good enough. Here again, we use advanced AI-based techniques to accurately classify feature-opinion pairs.

We classified close to 1000 opinion words and matched them to each feature. The infographic below shows groups of similar words that the AI algorithm clustered into a single feature, and the top positive and negative sentiments expressed by the Twitterati for each feature.

BĀHUBALI
SENTIMENT
ANALYSIS

	EQUIVALENT TERMS	POSITIVE SENTIMENTS	NEGATIVE SENTIMENTS
BĀHUBALI TEAM	TEAMS, TEAM, CREW	Brilliance, Superb, Grandeur, Amazing, Magnificient, Creative, Kudos	Hated, Sad, Fails, Gross
ACTING	ACTING, ACT	Perfect, Fantastic, Loved, Marvellous, Nice	Unsatisfactory, Not Real
VISUAL FX	VFX, CGI, SCENE, VISUALS, VISUALIZATION	Creative, Enjoyed, Grandeur, Heroism, Greater	Poor, Unrealistic, Disappointing, Dull Combat, Sucks
SECOND HALF	SECOND HALF, AFTER INTERVAL, SECOND PORTION, PART AFTER INTERVAL	Enjoyed, Justice, Loved, Engaging	Boring, Disappointed, Dragged, Stealing Tiring
DIRECTOR / RAJAMOULI	RAJAMOULI, SS RAJAMOULI, DIRECTOR, RAJAAMOULI, RAJAAMOULLI	Impressed, Applauding, Superb, Pride, Magnificient, Award, Dedication	Fool, Unrealistic, Crap, Failed

DATAWEAVE

While our technology can associate words with similar meaning, such as, 'part after interval' and 'second half', it can also identify spelling errors by identifying and grouping 'Rajamouli' and 'Raajamouli' as a single feature.

Adjectives like 'magnificent' and 'creative' were used to describe the Baahubali team positively, while words like 'boring', 'disappointed', and 'tiring' were used to describe the second half of the movie negatively.

Step 3: Sentiment Calculation

Lastly, we calculate the sentiment score, which is determined by the strength of the opinion-word, number of retweets and the time of tweet. A weighted average is normalized and we generate a score on a scale of 0% to 100%.

A Peephole into the Consumer's Mind

As more and more people express their views and opinions in the online world, there is more of an opportunity to use these data points to drive business strategies.

Consumer-focused brands use DataWeave's Sentiment Analysis solution as a key element of their product strategy, by reinforcing attributes with positive sentiments in reviews, and improving or eliminating attributes with negative sentiments in reviews.

Click here to find out more about the benefits of using DataWeave's Sentiment Analysis!

- Byom Kesh Jha
Data Scientist at DataWeave, 5th May, 2017

ARTIFICIAL INTELLIGENCE   BRAND PERCEPTION   BRANDS