Maximizing Conversion Value with Marketing Analytics and Machine Learning

A BrainTrust Insights Case Study

Executive Summary

The promise of digital marketing and the democratization of publishing tools was to put all brands on equal footing. The brands with the most skill at digital marketing and content creation would win, regardless of size. As advertising crept into the Internet, larger companies seized the advantage with big budgets and resources.

However, the introduction of machine learning technology is leveling the playing field once more. Bigger brands, plagued with inertia and aging infrastructure, are not able to adopt new technologies as quickly, creating an opening for more nimble companies to use machine learning to seize market share.

With machine learning, attribution analysis is more precise and nuanced; brands will understand the true impact of digital marketing channels, even with dozens or hundreds of steps leading to a conversion. Once a brand understands what truly drives business impact, it can extend its advantage using machine learning and predictive analytics to forecast likely business outcomes.

Predictive forecasts give us a starting point, a roadmap for planning that's truly data-driven. Instead of guessing when things are likely to happen in general, predictive analytics give us specificity, granularity to allocate budget and resources with precision. Nimble brands spend only when they need to, only when they are likely to generate outsized returns.

Finally, predictive analytics and machine learning help us to understand what will be on the minds of our customers and when, allowing us to create strategies, tactics, and plans of execution that astonish and delight customers. These technologies create magical customer experiences; brands seem to know exactly what we want during the exact time we're thinking most about it.

Situation Analysis

The role of content marketing is to attract interested audiences, engage them, and convince them to convert by demonstrating expertise. Measuring the impact of content marketing must be done regularly and frequently to assess the role and impact it plays in marketing effectiveness.

Measuring the impact of content marketing typically requires expensive attribution software such as Google Attribution 360™, which is a substantial financial burden for non-enterprise companies. Tools like Attribution 360™ use advanced machine learning techniques to determine the impact of each channel based not only on superficial data such as last touch, but also deep interactions among channels.

Once you complete attribution analysis, an enterprise marketer’s next step is to forecast, based on the attribution, what channels to allocate resources to, and use predictive analytics to determine when those resources will deliver maximum impact. Even enterprise tools like Attribution 360™ are unable to tie their insights to predictive forecasts, so marketers must find other alternatives to plan ahead.

Impact

The original promise of digital marketing was that the best content will win, regardless of who creates it. Digital marketing was supposed to level the playing field and permit small companies the same access and opportunities as enterprise companies.

In the absence of top-of-the-line attribution software, many non-enterprise companies do not have the ability to perform comprehensive attribution analysis, nor perform forecasting based on the attribution in order to maximize content marketing's impact. This leaves non-enterprise companies at a substantial disadvantage compared to their enterprise competitors.

We see this play out in content marketers' reasons for not developing a robust content marketing strategy, as documented in the 2018 Content Marketing Benchmarks by CMI/MarketingProfs:

Figure 1. Reasons why companies do not develop a content marketing strategy.

Small teams and lack of time are both based in resource utilization; enterprise companies are better-resourced to provide the necessary infrastructure such as machine learning-powered analytics.

We see the same when it comes to ROI and analytics for content marketing:

Figure 2: Why don't organizations measure ROI of content marketing?

The reasons given for why organizations do not have effective analytics in place range from poor governance to lack of knowledge/lack of resources.

What could small and mid-size companies do to close the analytics gap and create a more level playing field in content marketing? Could these companies use technology to partially make up for human-based resource shortages?

Implementation: Solving With Machine Learning

The solution to the resource gap in content marketing analytics is machine learning. With a combination of technologies, companies of any size - but especially small and mid-size companies - can partially close the performance gap of content marketing against their well-funded, well-staffed enterprise competitors. To close this gap, companies must leverage two classes of machine learning.

Machine Learning Attribution Analysis

The gold standard method for attribution and proving impact with machine learning is a technique called Markov chain attribution modeling. With Markov chains, every known conversion with more than one step is analyzed to determine the statistical frequency of specific paths to conversion. Once all the paths to conversion are analyzed and cataloged, machine learning software creates thousands or millions of iterative substitutions, then measures those substitutions against the known data.

This is akin to the board game Jenga, in which blocks are removed at random from a tower until the tower collapses. Much in the same way, channels are removed at random from conversion paths until the conversion no longer matches known models; by repeating this process many times, the relative importance of any given marketing channel is identified.

Without access to Google Attribution 360™, BrainTrust Insights set out to determine what drives conversions using machine learning and the attribution data available through a standard Google Analytics subscription - spinsucks.com in this instance. We extracted a rolling 365-day window of Goal Attribution data. Using the Markov Chain model that BrainTrust Insights built out using R , we ran the Google Analytics data through the algorithm. The results showed how channels were performing and wherein the customer journey they fell.

Time Series Forecasting / Predictive Analytics

Once the importance of each channel is determined with machine learning, the next step in the process is to determine how those channels will likely perform in the future. Using proven algorithms such as S-ARIMA, companies can forecast weeks or months in advance the likely performance of any given channel, then allocate resources appropriately to those channels.

Using the Google Analytics data from Spin Sucks, we forecasted each of the major digital channels' likely performance in the next 30 days as well as the next 52 weeks to help the site plan by media channel.

Once we ascertained the key driving channels - search being one - we turned our predictive capabilities towards the top search terms that Spin Sucks seeks ranking for in search engines. Using Google search historical data, we forecast the times those terms would be of greatest interest to the audience, then helped build a calendar by week of what topics Spin Sucks should be publishing to maximize audience interest.

Analysis and Insights

Attribution Analysis

We first look at the machine learning-powered attribution analysis. Direct/none leads the pack, but as is the case with so many websites, this is more a function of technological considerations than true marketing insight; direct/none is what Google Analytics assigns as a source when it doesn't receive data properly or doesn't recognize a data source.

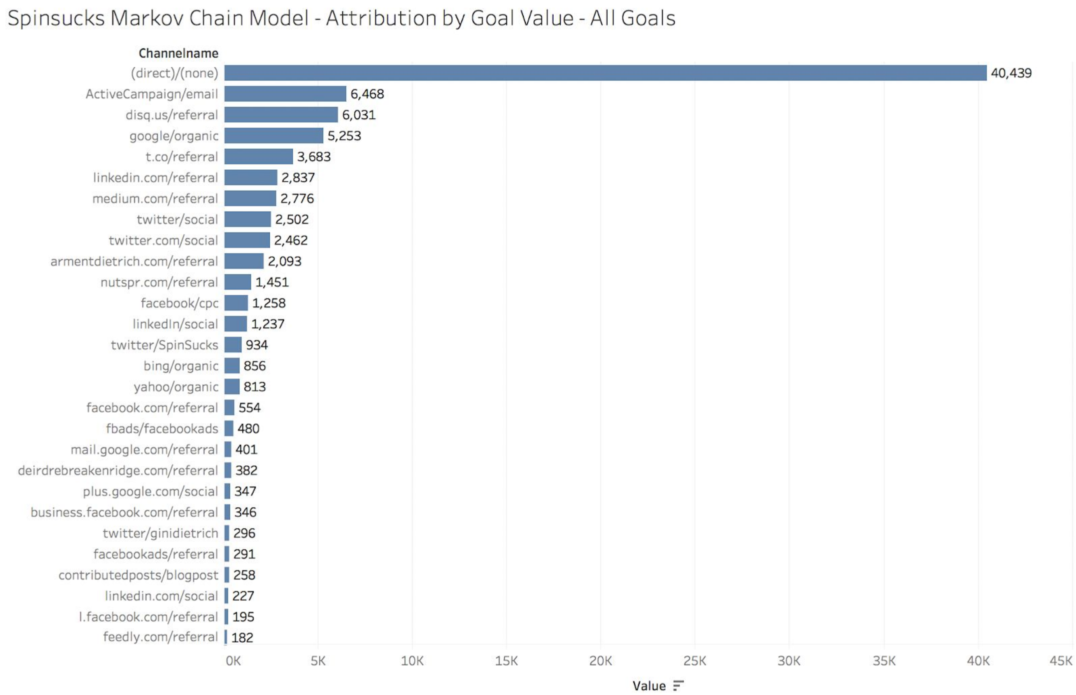

Figure 3: Markov Chain Models, All Channels

We remove Direct/none from the analysis for two reasons. First, we can't know the reasons why it's so high from the data itself (that requires separate technology audits), and second, Direct/none is not actionable. There's no action to take, no strategy to set from it.

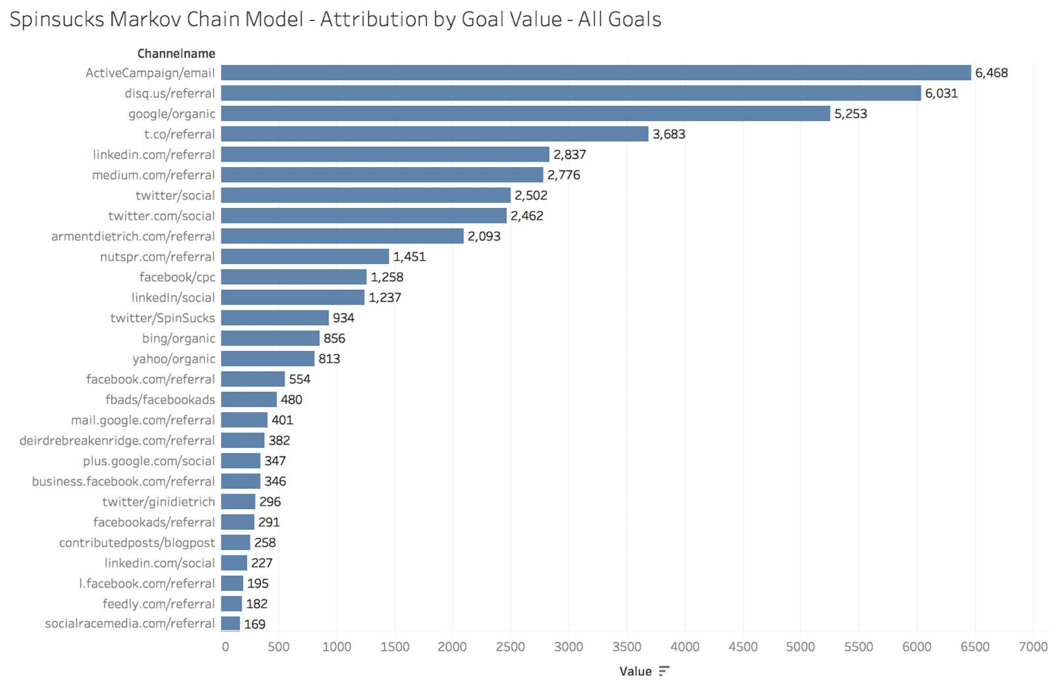

Figure 4: Markov Chain Models, Excluding Direct

When we look at the results without the Direct/none interference, we see a classic power law curve: a few channels account for many of the results, and many channels account for the rest. Leading the pack are email, referral traffic from comments, and organic search from Google, with Twitter traffic a distant fourth place.

What makes this type of attribution modeling useful and better than the out-of-the-box models in Google Analytics is that it takes into account weak or hidden effects much better. Weak interactions and hidden effects are when a channel participates in one step of multi-step conversions, but its contribution is overlooked by many attribution models.

For example, let's look at how the built-in tools attribute email in a last-touch model and a time-decay model (generally considered to be the best of what's in the box):

MCF Channel Grouping	Last Interaction		Time Decay		% change in Conversions (from Last Interaction)
	Conversions	Conversion Value	Conversions	Conversion Value	Time Decay
1. Direct	35,402.00 (56.05%)	[illegible]	33,661.57 (53.29%)	[illegible]	-4.92%
2. Organic Search	11,036.00 (17.47%)	[illegible]	11,531.59 (18.26%)	[illegible]	4.49%
3. Social Network	5,657.00 (8.96%)	[illegible]	6,119.43 (9.69%)	[illegible]	8.17%
4. Referral	5,500.00 (8.71%)	[illegible]	5,771.67 (9.14%)	[illegible]	4.94%
5. Email	3,868.00 (6.12%)	[illegible]	4,279.02 (6.77%)	[illegible]	10.63%
6. (Other)	936.00 (1.48%)	[illegible]	1,002.46 (1.59%)	[illegible]	7.10%
7. Paid Search	766.00 (1.21%)	[illegible]	798.93 (1.26%)	[illegible]	4.30%

Figure 5: Google Analytics Attribution Model Comparisons

We see email is given fifth place credit in both models, though the time decay model says that email should deserve some more credit. Neither of these out-of-the-box models show email's true dominance in helping make conversions happen.

What would happen if we made strategic decisions based on the out-of-the-box models? We might under-resource email marketing and over-resource other channels - and in turn, drive lower impact results.

Channel Forecasting

Once we know what channels to prioritize, we ran a Google Analytics predictive model through R using the same 365-day rolling window to see what the future holds per channel. We know, based on the attribution model, that driving email traffic, referral traffic, and organic search traffic are our priorities for the upcoming year.

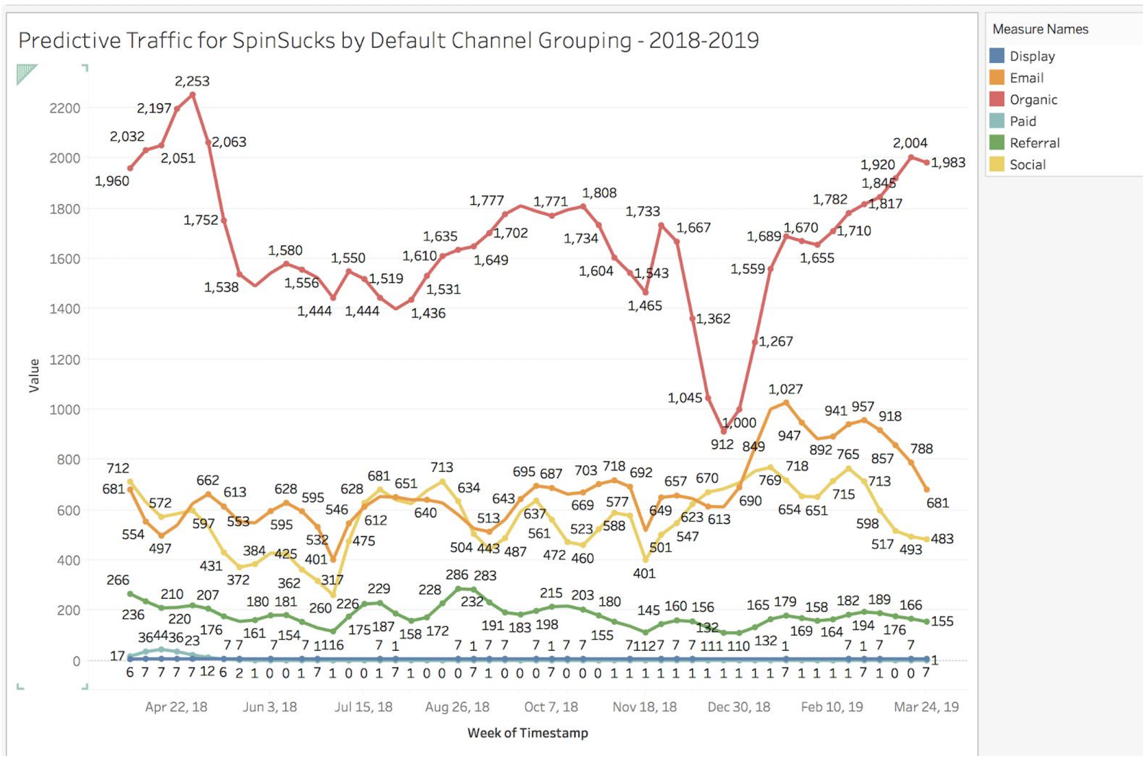

Figure 6: Predictive Forecast, By Google Analytics Default Channel Grouping

Above, we see the top channels - organic search in red, email in orange, and referral traffic in green, forecast for the next year. As with all predictive analytics, we focus on the peaks and dips. When a channel is forecasted to dip, we pull back some resources - staff, content, budget - and use the time to plan or bank new marketing materials for that channel. When a channel is forecasted to peak, we double down on it. On any given week, we may be pushing on one channel or another; some weeks (such as the week of January 7, 2019 above), we will be firing on all cylinders.

We see social media in the mix above as well, as the yellow line, but because social media didn't fit prominently in our attribution model, we won't use it as a priority channel. Certainly, a digital marketing strategy focused on email, referral traffic, and organic search will have a halo effect on what we share on social media, but our attribution models indicate a social-first strategy would not be the way to go.

Campaign Forecasting

We've established the importance of each channel and the likely traffic to Spin Sucks with predictive analytics; our final step in the process is to forecast what thematic campaigns we should focus on. Using historical search data from Google Trends, we built a predictive forecast model for the key terms relevant to Spin Sucks' business and content strategy: marketing courses, media training, PR courses, PR training, public relations courses, social media courses, social media training.

Campaign Forecast

	Group						
Week of Ds	marketing-courses	media-training	pr-courses	pr-training	public-relations-c..	social-media-cour..	social-media-train..
April 8, 2018	80.95		65.10	82.06	78.82	58.06	59.37
April 15, 20..	59.23		68.73	72.79			61.44
April 22, 20..	50.09	67.88	59.22	54.61			66.72
April 29, 20..	50.59	88.32			70.29		66.30
May 6, 2018		76.72			100.00		57.52
May 13, 20..					66.48		
May 20, 20..							51.63
May 27, 20..	50.48						66.07
June 3, 2018	61.05	56.70					83.30
June 10, 20..	52.44	86.68		64.49			88.64
June 17, 20..		84.59		58.05	64.81		78.60
June 24, 20..		58.46					63.31
July 1, 2018							51.37
July 15, 2018	57.85	65.70					
July 22, 2018	60.21	90.12		54.20		60.36	
July 29, 2018		100.00			52.41	67.35	64.64
August 5, 2..		93.25				63.73	87.49
August 12, ..		79.54				60.38	99.29
August 19, ..		67.66			54.93	61.71	87.63
August 26, ..	52.13	62.51			87.74	61.67	66.05
September ..	51.57	66.36		74.38	92.34	55.17	55.95
September ..	53.95	76.72		85.38	62.27		65.25
September ..	62.93	84.07		68.84			86.55
September ..	72.30	79.19		58.89		56.87	100.00
September ..	75.91	66.73		69.19	51.32	64.05	94.22

Figure 7: Predictive Content Calendar by Week

Using our predictive forecasting, we develop a calendar for each week of the year ahead. When a cell is colored in, we know that particular topic will be a focus point that week. When a cell is empty, that week isn't a focus point for that topic. Above, we see that for the next 6 months, social media training will be a hot topic for most of the time; PR courses will largely be absent as a topic after April 2018.

With these insights by week, Spin Sucks can now calibrate a number of different activities to the calendar, such as:

- Public relations campaigns timed to land coverage during specific weeks, based on publication editorial calendars
- Inbound link campaigns timed to bring in links before critical times, to ensure search engines are finding the site frequently
- Referral traffic campaigns such as guest blogging, to attract traffic when audiences are thinking about the above topics most frequently
- Email campaigns timed to arrive in inboxes exactly when audiences will be thinking about a given topic the most

We align these with the projected traffic patterns from Google Analytics™ to further inform our overall strategy and resource allocation. When Google Analytics™ traffic and audience interests overlap, we know to significantly increase our investments.

Summary

While much is made of “all-in-one” marketing analytics solutions, the reality is that best-of-breed solutions often require multiple, disparate components. In our walkthrough of Spin Sucks’ needs, we built on a strong foundation of Google Analytics™ data, but enhanced it with machine learning attribution analysis.

Once we determined what was truly influencing business results, we doubled down on machine learning with predictive analytics to set overall resource allocation and planning.

Finally, we turned our predictive capabilities towards our audience’s interests, providing granular insights about specific topics the audience wants, aligning them to our overall strategy and plan of execution.

By using machine learning, we move closer to fulfilling digital marketing’s promise, allowing smaller brands to punch above their weight. Instead of competing toe-to-toe with larger competitors, smaller brands can focus their resources on the most likely successful outcomes and beat bigger competitors with precise timing and spending.

Contact Information

Katie Lioy
Co-founder and CEO
klioy@braintrustinsights.com
https://BrainTrustInsights.com
@TrustInsights