

IPG MEDIABRANDS



AD VISION 2.0



Using Machine Learning Power of Google's Cloud Vision API



EXECUTIVE SUMMARY

Introduction

In this follow up research on www.mediabrandsadvison.nl, we again used the machine learning technology of Google's Cloud Vision API to test the impact of creatives on ad effectiveness. We tested over 13.000 online ads from over 50 product categories. We ran these ads through the Vision API to find out what ad characteristics, labeled by the Vision API, drive the most success in terms of CTR. In this follow up, we test new variables, use different modeling techniques, and create an ad upload possibility on our new website to estimate the CTR for your banner. Because of our first award winning project, which was the first ad effectiveness study using Machine Learning as core methodology in the Netherlands, we could use most of our previous learnings to improve the quality of this research.

Methodology

To collect all ads and their performance metrics from our ad servers we used the DCM/DFA Reporting and Trafficking API, and Java client libraries. We loaded all creatives for which performance metrics were still available, and stored them specified per day. We uploaded all images into Google Cloud Vision API requesting Label, Text, Landmark, Logo Detection, Face Detection and Object Detection, and stored the results for analysis. Google provided funding for this research.

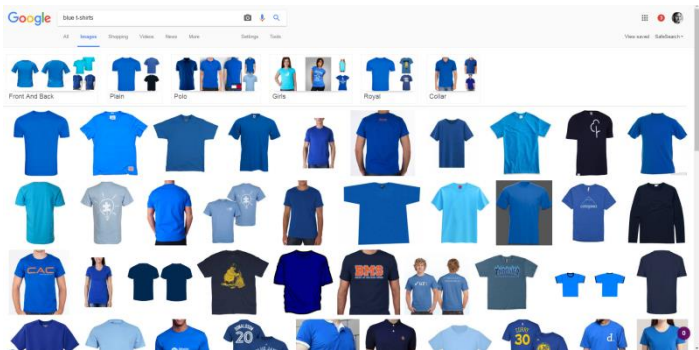
Most important findings

- ❑ Large ads perform better than small ads
- ❑ Horizontal is the best shape compared to vertical
- ❑ Dynamic ads perform better than static ads. However, for large ads the differences are small.
- ❑ Small moving ads perform much better than small static ads
- ❑ Including a logo has a positive effect on the CTR
- ❑ Including an object increases the CTR
- ❑ Adding a face or person does not always has a significant effect on the CTR
- ❑ Using primary or secondary colors instead of mostly white, black or gray increases the CTR

AD EFFECTIVENESS USING MACHINE LEARNING TECHNIQUES

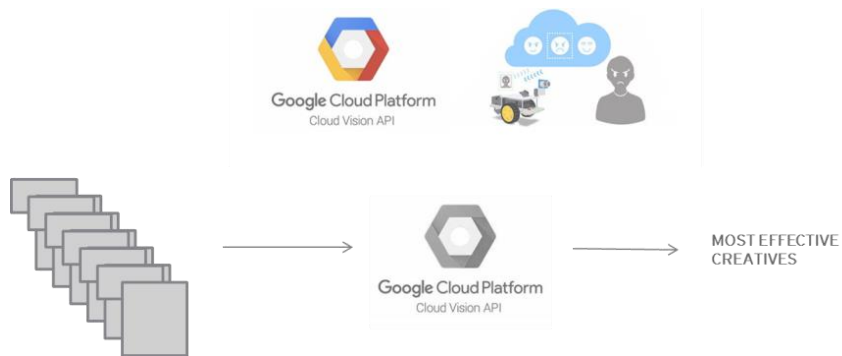
Google Images Search

When you use Google Search images, and you search for 'blue t-shirts', one will find images with blue t-shirts in it, even though they are not specifically tagged with 'blue' and 't-shirt'. The technology used by Google is trained to recognize visual characteristics of images, using machine learning. The machine recognizes color, logos, texts, fonts, faces, animals, et cetera. It also recognizes context and can, for instance, associate pink with 'female' in a fitting context.



Google Cloud Vision API

We used this technology to test the effectiveness and characteristics of thousands of online display and animated ads. Moreover, we studied the impact of these characteristics on the click through rate (CTR) for online advertisement. Furthermore, we designed a new methodology based on a k-means cluster algorithm and the LAB color space to determine dominant colors in an ad.



LABELING ASSOCIATIONS & RECOGNIZING OBJECTS

From object recognition to facial expressions

It is definitely interesting to see how “the machine” works and what associations it makes. It not only measures the number of letters, pixels, RGB colors, fonts and logos, but also recognizes animals, famous buildings like the Eiffel Tower, objects like a sail boat or a car, and facial emotions of people like joy, sorrow and anger. Moreover, Cloud Vision API associates colors and visuals. An example from our analysis: An ad with a dominant pink color and a heart shape in it was not only labeled ‘pink’ and ‘heart’ but also as ‘female’.

Most of the time the application is right, like when it recognizes that the woman wears swimwear. However, sometimes we see some strange associations, like with the mobile phone which he wrongly recognizes in the objects.

The advertisement shows a woman on a yellow inflatable in a pool, a mobile phone, and various household items. The text in the ad is: "Alleen op plus. Alleen op plus; gratis bezorging is bezorging bij bestelling van producten. gratis bezorging is bezorging bij bestelling van deze producten. gratis bezorging is bezorging bij bestelling van deze producten." The Plus logo is repeated at the bottom of the ad.

LABELS	OBJECTS	LOGOS
<ul style="list-style-type: none"> yellow swimwear sun tanning summer undergarment vacation inflatable leisure advertising happiness product 	<ul style="list-style-type: none"> muscle thigh product technology water plastic brand font Bottle multimedia Electronic device junk 	<ul style="list-style-type: none"> food liquid blue green aqua turquoise azure organism sky computer wallpaper turquoise pattern
	<ul style="list-style-type: none"> Person Swimwear Bottle Mobile Phone Kitchenware Tableware Food Plate Shorts Tin Can Box 	<ul style="list-style-type: none"> Plus Retail FACE FALSE

METHODOLOGY

Getting all images from DCM

To collect all ads together with their performance metrics from our ad servers, we used the DCM/DFA Reporting and Trafficking API and Java client libraries. We imported all creatives for which performance metrics were still available, and stored them specified per day. For animated ads, the URL's where reconstructed and then each of the ads was automatically opened for 30 seconds using PhantomJs or SlimerJs. During these 30 seconds, screen captures where taken 10 times each second, giving us up to 300 frames for each ad.

Getting key frames from animations

Approximately 15% of our analyzed ads are static ads, and 85% dynamic. For the dynamic ads, we created a python script that is able to analyses dynamic animations based on the frame to frame changes. We collected up to 300 frames ("snapshots") for each ad. Next, we collected the unique frames for each ad in order to identify the different characteristics. From this we also acquired properties of the animation, like animation length, the number of loops and the type of animation (continuous or transitions).

Running Cloud Vision API

After creating the multiple keyframes, we uploaded all images into Google Cloud Vision API requesting Label, Text, Landmark, Logo Detection, Face Detection and object detection, and stored the results as lists in a Mongo Database. A second Python script determines the 7 pixel clusters, explained on next page. Furthermore, it calculates the luminance and saturation for each ad. Finally, it calculates size, positions (logo, object) from the presented coordinates by Google's Vision API and made selections based on confidence scores (logo, object, face, label).



Moving ad, created 5 unique shots, results stored as being one ad.

NEW METHODOLOGY FOR DOMINANT COLORS

K-Means clustering algorithm

Clustering is a method to determine groups of objects. K-means clustering treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. In terms of ads, the cluster algorithm determines clusters within the pixels of an ad. We allowed the algorithm to use 7 clusters for each ad and then calculated the coordinates of their centers using the RGB color space. A RGB color space can be understood by thinking of it as all possible colors that can be made from three colored lights for red, green, and blue.

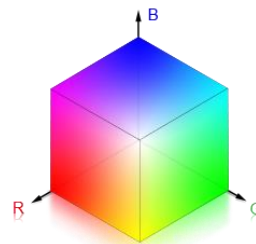
Using the LAB color space

We used this LAB color space because it is designed to perceptually uniform with respect to human color vision, meaning that the same amount of numerical change in these values corresponds to about the same amount of visually perceived change.

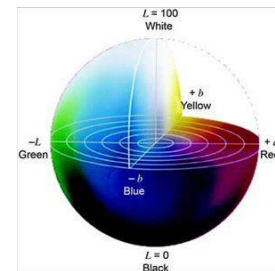
<https://blog.xkcd.com/2010/05/03/color-survey-results/>

After calculating the RGB coordinates for each cluster, we were able to assign a LAB color to each cluster. For this end, we used the LAB color space from the XKCD research (A research where over five million colors were named across 222,500 user sessions) to create a list of color values in the 'LAB' color space with corresponding color labels.

Following the results, this approach for color recognition is the most accurate one. Furthermore, applying this methodology enables us to assign up to seven dominant colors to each ad, in contrast to the one dominant color for each ad in the previous project. Each color recognized in an ad is presented together with a color fraction which is used as a threshold for being dominant yes or no.



RGB color space



LAB color space

IPG MEDIABRANDS

CTR

Analyzing the Click Through Rate



CLICK THROUGH RATE

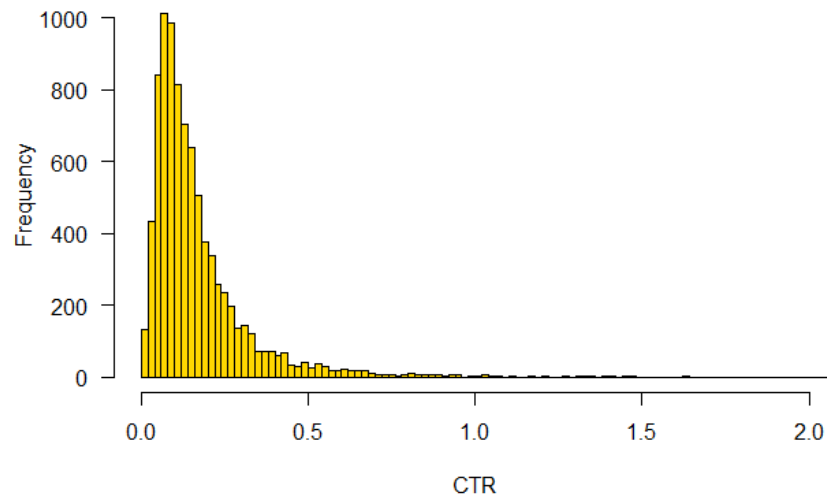
CTR

We tested over 13.000 ads and ran them through Cloud Vision API. Only ads with at least one click and more then 5.000 impressions were included. Furthermore, we excluded ads having an unrealistic viewable CTR. For dynamic formats we created a script that splits the ad into frames.

The average CTR over all ads is 0,17% with a median of 0.12%. The max CTR in our sample is 2.33%. Furthermore, we analysed that 95.5% of the ads in our sample achieved a CTR between 0% and 0.5%.

	CTR (%)
Min	0.000051
Max	2.333436
Mean	0.17435
Median	0.12413

Click Through Rate



IPG MEDIABRANDS

MODEL

Generalized linear model





MODEL USED

Do we create separate models for static and dynamic ads?

There have been several researches on whether static or dynamic banners perform better in terms of CTR. While most people think that dynamic banners perform better as they offer an opportunity to communicate multiple messages, research shows that actually static banners perform better. This is due to the fact that dynamic banners are usually experienced as visually distracting and slow. Therefore, we expect that the variables which we are including in our model will show different effects across the dynamic types. However following our analyses, using one model which includes a variable indicating whether or not a ad is dynamic, outperforms two separate models for static ad and dynamic ads. Therefore, following our conclusions based on statistical analyses, we created one model for estimating the CTR.

Variables tested to estimate the CTR

To estimate CTR we used data obtained by three sources. We collected data regarding clicks, viewability and impressions from DCM. Secondly, information about the presence of objects, labels, faces, logo's and their characteristics (position, size and scores) were obtained by Google's Vision API. Finally, Marketing Sciences calculated shape, sizes, colors, saturation and luminance.

Generalized Linear Model (GLM)

The CTR is not a continuous variable (because it is restricted for values < 0 and > 100) and moreover, it is not perfectly distributed between 0-100 percent (but mostly between 0 and 0.5 percent). Therefore, we were not able to use the standard linear regression.

Following statistical assumptions, we used a generalized linear model together with the quasi-binomial family. The generalized linear model allows response variables (the variables used to estimate/explain CTR) to have an error distribution other than a normal distribution. In our case, we used the quasi-binomial family of the GLM. Using this family, we were able to describe additional variance in the data that cannot be explained by a Binomial distribution alone. Moreover, the quasi-binomial family has proven to perform well in situations where one aims to model a process which can be described in terms of failures and successes. In our case, we were able to describe the CTR in terms of impressions - clicks (failures) and clicks (successes).

RETAIL

Retail

IPG MediaBrands serves a lot of clients who are active in the Retail Industry. Actually 36% of our ads tested belongs to Retail companies. Therefore, we splitted our dataset in Retail and non Retail ads. For this end we categorised our clients in the dataset in Retail and non Retail clients.

We analysed a minimum difference in average CTR in favor of the Retail companies. However, during the modeling process, we could not find any statistical prove of the effects of being a Retail company. This does not have to mean that there is no difference in the CTR of Retail companies in comparison with other industries, but can be caused by the variety of companies in our dataset.

On the right hand side one finds all the objects recognized by the Vision API related to food and consumer goods. As one can see, the Vision API recognizes a lot of Retail related products.

	%	CTR (%)
Retail	36.1	0.175
Non Retail	63.9	0.174

apple	pumpkin	alarm clock	flowerpot	pillow
asparagus	salad	axe	food processor	plate
baked goods	sandwich	backpack	football	punching bag
banana	snack	ball	football helmet	racket
bell pepper	squash	baseball bat	fork	refrigerator
bread	strawberry	bathroom cabinet	frying pan	remote control
broccoli	sushi	bed	furniture	rugby ball
cabbage	taco	bench	glasses	ruler
cake	tin can	bicycle helmet	guitar	scissors
candy	tomato	bookcase	handbag	shower
carrot	vegetable	bowl	headphones	spoon
cheese	waffle	box	helmet	stapler
cookie	Watermelon	cabinetry	home appliance	stool
dessert		camera	infant bed	suitcase
egg		candle	kitchen appliance	sunglasses
food		chair	kitchenware	surfboard
french fries		chest of drawers	kite	table
fruit		coffee cup	lamp	table tennis racket
grape		coffee table	lantern	tablet computer
grapefruit		coffeemaker	laptop	tableware
hamburger		computer	light bulb	teddy bear
hot dog		keyboard	lighting	television
ice cream		lipstick	computer monitor	tennis ball
lemon		loudspeaker	corded phone	tennis racket
mango		loveseat	couch	tent
mushroom		luggage & bags	crown	toilet paper
orange		microphone	dairy	toy
pancake		microwave oven	desk	treadmill
pasta		mirror	dishwasher	umbrella
pastry		mixer	doll	vase
peach		mobile phone	drill	volleyball
pear		musical instrument	drum	washing machine
pineapple		musical keyboard	dumbbell	watch
pizza		nightstand	earrings	whisk
popcorn		pen	filing cabinet	window blind
potato		flag	perfume	

IPG MEDIABRANDS

SIZE

The impact of size on ad effectiveness



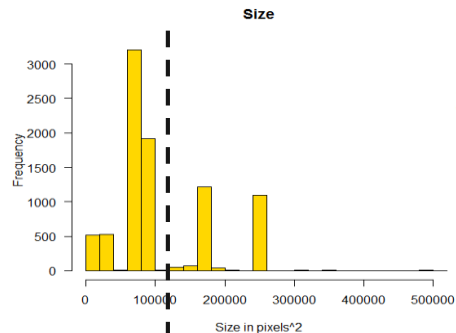
THE IMPACT OF AD SIZE

The impact of size

Most of the ads tested, were in the same size range (50.000 - 100.000 pixels²), with an average CTR of 0,1337%. Inspecting the graph, one can see different size ranges with peaks in number of ads. For this reason it might be interesting and helpful to create different size groups while modeling CTR. If we split the data into 4 groups reflecting the four peaks in the graph, the last group (around 250.000 pixels²) has the highest average CTR.

Creating two categories based on size

For modeling purposes, we splitted the ad size into two categories. Using R, we created a cluster algorithm to determine the optimal number of groups to be splitted and to determine the boundaries for these groups. The cluster algorithm determined to split the dataset in two groups at 128.000 (pixels²). The small group contains 71.7 % of the ads in our dataset with area's below 128.000 squared pixels. We analysed that larger ads tend to have a higher average CTR, as we expected.



IAB SIZES

320	50	16000
320	100	32000
738	90	66420
120	600	72000
300	250	75000
320	240	76800
336	280	94080
160	600	96000
320	480	153600
300	600	180000
970	250	242500

Min	10,800
Max	1,800,000
Mean	111,446
Median	94,080

Shape	Area 1 - Small 0 - 128.000	Area 2 - Large 128.000 - 1.800.000
Min	10,800	145,600
Max	128,000	1,800,000
Mean	71,329	212,901
Median	75,000	180,000
Fraction (%)	71.7	28.3
CTR (%)	0.145	0.249

THE IMPACT OF AD'S MOVING AND THEIR SHAPE

The impact of shape

The shape of the ad has a serious impact on its CTR. Horizontal ads (70 percent of our dataset) are more effective (overall average CTR of 0,1842) than vertical and squared ones. Moreover, horizontal and large ads achieved the highest average CTR, while vertical and small ads achieved an average CTR of just 0.09 percent.

However, we have to be mindful that only 0,5 percent of our ads is squared in the data set, hence we did not included squared ads in the model.

Shape	Fraction Overall (%)	CTR Overall	CTR - Small 0 - 128.000	CTR - Large 128.000 - 1.800.000
Vertical	29.5	0.151	0.09	0.215
Horizontal	70.0	0.184	0.16	0.286
Squared	0.05	0.132	0.13	0.201

Dynamic and Static

To be able to analyse moving/dynamic ads, we took snapshots every 10 milliseconds using scripts created by our Tech Department. Next, we removed all identical shots and combined the remaining unique shots. From this, we analysed each unique shot, and combined the results. Finally, we stored the data in a dataset for analyzing and modeling purposes. Totally, 86.8 % of the data set contains dynamic ads.

While most people think that dynamic banners perform better as they offer an opportunity to communicate multiple messages, some researches show that actually static banners perform better. This is due to the fact that dynamic banners are usually experienced as visually distracting and slow. However, in terms of CTR, we analysed that dynamic ads do have a better average performance than static ads. Nevertheless, for large ads, the average CTR of static and dynamic ads have an equal average CTR.

	CTR Overall	CTR - Small 0 - 128.000	CTR - Large 128.000 - 1.800.000
Static	0.154	0.118	0.249
Transitions	0.178	0.148	0.254
Continuous	0.176	0.153	0.226
Dynamic (continuous + transitions)	0.177	0.149	0.249

IPG MEDIABRANDS

COLOR

The impact of color on ad effectiveness

THE IMPACT OF DOMINANT AD COLOR

Color

The effects of colors in marketing and media has been frequently analysed during the last years. Color attracts attention, impacts trust, increases recognition, and associates emotion. Therefore, we believe it could be very interesting to examine the effects of colors used in online ads. For this end, we created a new methodology to determine dominant colors in ads to estimate the effect of a single color on CTR.

Keep in mind that there is no such thing as a perfect color. Actually, emotions regarding colors can be different for people from different countries, cultures or ages. Moreover, research has shown that color choices perceived to be appropriate to a product/brand are a significant factor for consumer buy-in.

Because it is extremely hard to add all different colours separately in the model, we made three colour groups. First we created the primary colors: red, yellow and blue. Second, we combined the secondary colors purple, green and orange into one group. Finally, we combined black, white and gray. As a result, we left brown and pink out of the model.

Color	Fraction Yes (%)	CTR yes (%)	Fraction No (%)	CTR No (%)
Blue (primary)	18.7	0.179	81.3	0.173
Red (primary)	31.2	0.178	68.8	0.173
Green (secondary)	6.6	0.164	93.4	0.175
Black (no color)	9.7	0.186	90.3	0.173
White (no color)	16.7	0.178	83.3	0.174
Pink	5.4	0.172	94.6	0.174
Yellow (primary)	5.9	0.177	94.1	0.174
Purple (secondary)	0.3	0.224	99.7	0.174
Brown	0.4	0.158	99.6	0.174
Gray (no color)	2.8	0.171	97.2	0.174
Orange (secondary)	2	0.201	98	0.174

We analyzed that adding primary or secondary colours (or both) can increase your CTR, while having black, white or gray as only dominant colours will decrease your possibilities of obtaining an high CTR. This result satisfies the findings in other studies which claiming that colorfulness can increase your CTR.

LUMINANCE

Luminance

In this study we also examined the effects of luminance(brightness) and saturation. Luminance refers to how much white (or black) is mixed in the color while Saturation indicates the amount of grey in a color. A saturation value of 0 indicates mostly grey while 100% luminosity is white.

Luminance describes the measurement of the amount of light emitting, passing through or reflected from a particular surface from a solid angle. It also indicates how much luminous power can be perceived by the human eye. This means that luminance indicates the brightness of light emitted or reflected off of a surface. In the display industry, luminance is used to quantify the brightness of displays.

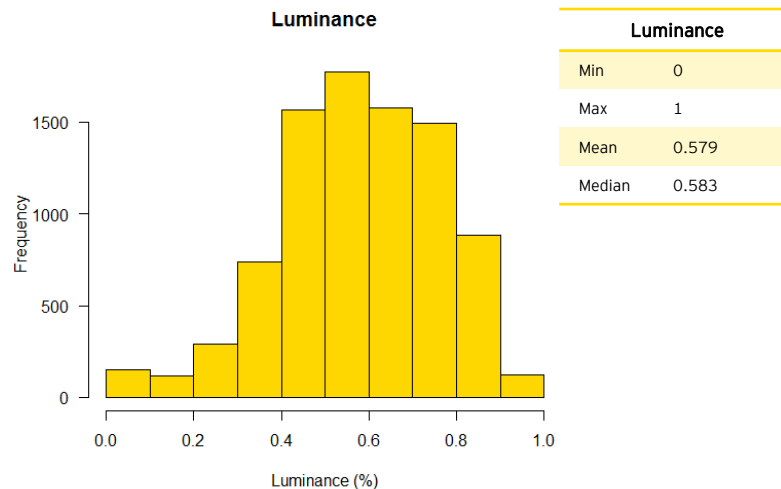
In research regarding the effect of luminance: *"Is Brighter Always Better? The Effects of Display and Ambient Luminance on Preferences for Digital Signage"* the authors concluded the following:

"Our results show that the mantra "the brighter the better" is not always true. There appears to be a point at which increasing image luminance has no benefit for, and may produce a decrement in, viewer satisfaction."

<https://pdfs.semanticscholar.org/ff63/39c90b3b45053f0518823c765bd1f4976df1.pdf>

For this reason we examined the effects of luminance and aimed to find a optimal region for the amount of luminance in an ad.

However, we could not find a statistical prove of the effect of luminance and therefore did not include luminance in our models.



SATURATION

Saturation

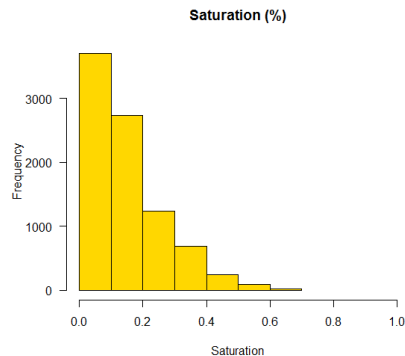
Color saturation refers to the intensity of color in an image. The term hue refers to the color of the image itself, while saturation describes the intensity (purity) of that hue. When color is fully saturated, the color is considered in purest (truest) version. Primary colors red, blue and yellow are considered truest version color as they are fully saturated.

When the saturation is zero, what you will see is a shade of gray. So, saturation refers to how strong or weak a color is (high saturation being strong).

The saturation value has recently been found to have several interesting effects on other perceptions. This was discovered by Hagtvedt & Brasel (2017), who conducted multiple experiments to test the effects of color saturation. In one of these experiments, the test subjects had to guess the size of two similar objects presented simultaneously, only differing in color saturation. More than 65% of the subjects judged the object with a higher color saturation to be larger than the size of its lower saturated counterpart.

The researchers also found that the color saturation of an object does not only influence the size perceptions of the object itself, but also of its direct surroundings. When they showed test subjects a picture with an object in a room, the height of that room was judged to be significantly lower when the object's color saturation was high.

Following their findings, we analysed the effect of saturation on CTR. However, it is difficult to interpret the saturation on moving ads. This is due to the many snapshots we take, hence the overall saturation seems to be an unreliable variable for our purposes.



Saturation	
Min	0
Max	0.679
Mean	0.15
Median	0.116

<https://www.newneuromarketing.com/the-size-of-colors-intenser-means-bigger:>

IPG MEDIABRANDS

LOGO | FACE | OBJECT

The impact of logo, face and object on ad effectiveness



GOOGLE VISION API – LOGO, FACE, OBJECT

Variables obtained using Google's Vision API

Information regarding the presence of a logo, object or face are obtained by using Google's vision API. Keep in mind that, although the API has improved since the previous project, there are still multiple examples of wrongly identified characteristics. Therefore, we decided not to include information regarding characteristics of logo, face and object in our model.

Following the results of the Vision API, 76,28% of the ads contain a logo. Ads containing a logo have a higher average CTR than ads without a logo. Furthermore, some ads contain a face (6.2 %), which also results in a higher average CTR. Finally, ads containing one or more objects do have a higher average CTR. However, the differences for all three variables are relatively small.



Face	%	CTR (%)
Yes	6.2	0.179
No	93.8	0.174

Logo	%	CTR (%)
Yes	76.28	0.175
No	23.72	0.173

Object	%	CTR (%)
Yes	63.96	0.175
No	26.04	0.172

IPG MEDIABRANDS

TEXT

The impact of text



THE IMPACT OF THE USE OF TEXT

Unique words

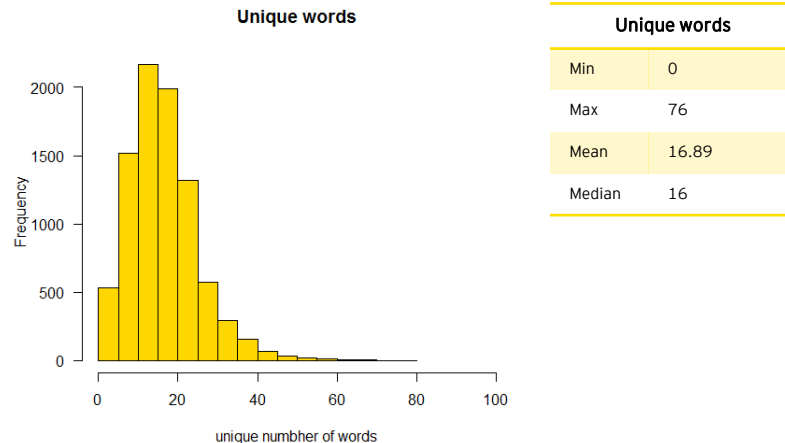
For the analysis of text, we determined the unique words in an ad. By analyzing unique words, we created a variable for measuring the amount of text used in an ad. The initial assumption to use this variable was that zero text could decrease CTR, while too much text could also lead to a decrease in CTR. Therefore, we aimed to determine an optimum for the amount of text used. We defined zero text for ads having zero or one unique words.

Text

We analysed two types of text: Call to Actions and money related text. For call to actions we used words as: click, press, continue, download, win, look, check, find, discover, read, book, sign, subscribe, join and claim. For money related tekst we used the words: euro, money, discount, 2=1, 3=1, for sale, offer, cheap, safe and free.

Wrong tekst detection

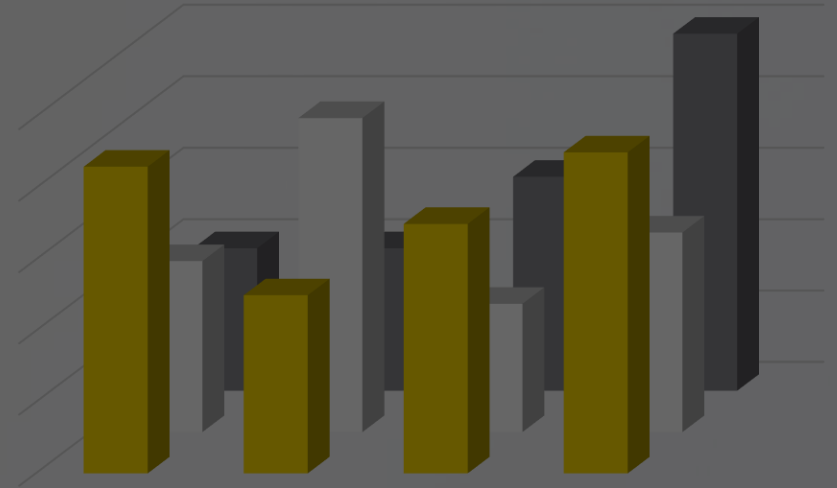
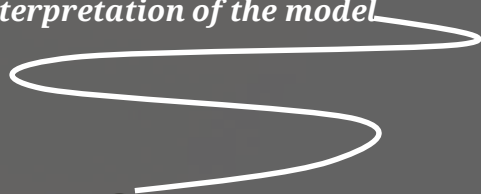
Analyzing the call to actions and money related words, we concluded that the API made too many mistakes in recognizing words in text. In detail, a lot of call to actions were not recognized and the API sometimes splitted single words in more words (money became "mon" and "ey"). As a result, we were not able to create reliable variables for estimating the effect of text on the CTR and therefore unfortunately, we could not include any text related variables in our model.



IPG MEDIABRANDS

RESULTS

Interpretation of the model



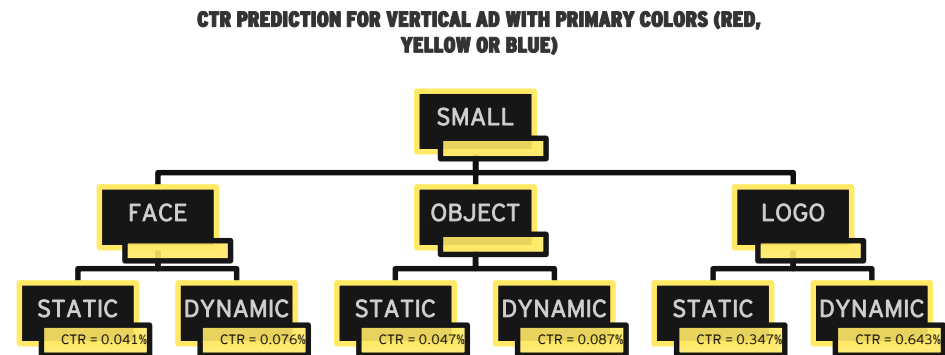
CTR PREDICTION COMPARISON

CTR Prediction

Using the estimates of our Generalized Linear Model, we are able to identify the effects of different characteristics of an ad. We now present an example to clarify this further:

Suppose we have a small vertical static ad with mainly primary colors, in which the logo is present, but there are no faces or objects detected. According to our model, we will predict a CTR of 0,347% for such a banner.

Changing the ad from static to dynamic will lead to a CTR prediction of 0.643%. If we also would include an object, and increase the size from small to large, the model calculates a predicted CTR of 1.302%. Adding a face in the banner would decrease the predicted CTR to 1.175%. This emphasizes the collaboration of our variables in the model, which in the end will give a good prediction of the CTR.



CTR PREDICTION EXAMPLES

CTR Prediction Examples

We splitted our dataset in a training and testing dataset, which means 80% of the datapoints (i.e. the ads with their corresponding characteristics) is used to estimate the model and 20% of the datapoints is used to determine the predictive accuracy of the model. Thus, we use the estimated coefficients to predict the CTR for the ads in the testing dataset and compared it to the actual CTR (which of course is known as these ads already generated impressions and clicks).

In the table we have shown the actual vs. predicted CTR for 10 randomly chosen ads in the testing dataset. It can be seen that the deviation from the actual CTR is very small, meaning the model has a high predictive accuracy.



Ad	Actual CTR	Predicted CTR by our model
1	0.118%	0.127%
2	0.183%	0.182%
3	0.141%	0.134%
4	0.122%	0.124%
5	0.202%	0.195%
6	0.144%	0.140%
7	0.123%	0.125%
8	0.158%	0.158%
9	0.079%	0.086%
10	0.144%	0.141%



CONCLUSIONS

Most important findings

- ❑ Large ads perform better than small ads
- ❑ Horizontal is the best shape compared to vertical
- ❑ Dynamic ads perform better than static ads. However, for large ads the differences are small.
- ❑ Small moving ads perform much better than small static ads
- ❑ Including a logo has a positive effect on the CTR
- ❑ Including an object increases the CTR
- ❑ Adding a face or person does not always have a significant effect on the CTR
- ❑ Using primary or secondary colors instead of mostly white, black or grey increases the CTR



IPG **MEDIABRANDS**

Mediabrand Marketing Sciences



More info: menno.van.der.steen@mbww.com

