VortexDNA Example Application

One Possible Application of VortexDNA within a Search Engine or Product Recommendation Environment

VortexDNA search and recommendation technology complements a client's existing technology. As such, it can be implemented in a variety of ways depending upon the needs and circumstances of the client. How might VortexDNA technology be used to optimise a search engine's or product recommendation's algorithms? One possible application of the technology is presented below.

Baseline Data Needed

We can track the percentage of times that users click on links in each link position in a set of search results or product recommendations, and can sort this data based on whether the link has a High Relevance Score or Low Relevance Score (as assessed by VortexDNA). An example table of this data would look as follows:

Table Showing Percentage of Clicks for Links of Each Type			
Link Position	High Relevance Score links	Low Relevance Score links	Baseline (all links)
1st	23.6%	17.6%	20.6%
2nd	20.9%	16.9%	18.9%
3rd	8.1%	5.1%	6.6%
4th	6.8%	4.8%	5.8%
5th	4.2%	2.6%	3.4%
6th	3.5%	2.3%	2.9%
7th	3.0%	2.0%	2.5%
8th	1.8%	1.2%	1.5%
9th	1.5%	1.1%	1.3%
10th	0.6%	0.4%	0.5%

Let us convert this to show the increased percentage of times that links are clicked on when they have a High Relevance Score and the decreased percentage for links with a Low Relevance Score, compared with the baseline percentage for that link position in the set of search results:

Table Showing Increased Percentage of Clicks Compared with Baseline			
Link Position	High Relevance Score links	Low Relevance Score links	Baseline (all links)
1st	3.0%	-3.0%	0.0%
2nd	2.0%	-2.0%	0.0%
3rd	1.5%	-1.5%	0.0%
4th	1.0%	-1.0%	0.0%
5th	0.8%	-0.8%	0.0%
6th	0.6%	-0.6%	0.0%
7th	0.5%	-0.5%	0.0%
8th	0.3%	-0.3%	0.0%
9th	0.2%	-0.2%	0.0%
10th	0.1%	-0.1%	0.0%

Optimisation Algorithm

Consider a particular set of search results. Without any further information about these links, we would make the assumption that, since the search engine or product recommendation algorithm has presented the results in this order, the implicit expected probability of each link being clicked is equal to the baseline for that link position. For example:

Table Showing Implicit Probability of Clicks for a Set of Search Results Based on Baseline Percentages		
Link Position	Link Name	Implicit Probability of Being Clicked
1st	Link A	20.6%
2nd	Link B	18.9%
3rd	Link C	6.6%
4th	Link D	5.8%
5th	Link E	3.4%
6th	Link F	2.9%
7th	Link G	2.5%
8th	Link H	1.5%
9th	Link I	1.3%
10th	Link J	0.5%

However, the expected probabilities, as we have seen above, are different when VortexDNA has rated a link as having a High or Low Relevance Score for the user. For example, if the 2nd, 4th, 7th and 9th links have High Relevance Scores for this user and the others have Low Relevance Scores, the corrected probabilities would be as shown below:

Table Showing Corrected Probabilities Based on Relevance Scores of Links					
Link Position	Link	Implicit Probability of Being Clicked	Relevance Score of Link for User	Percentage Change from Baseline	Corrected Probability of Being Clicked
1st	Link A	20.6%	Low	-3.0%	17.6%
2nd	Link B	18.9%	High	2.0%	20.9%
3rd	Link C	6.6%	Low	-1.5%	5.1%
4th	Link D	5.8%	High	1.0%	6.8%
5th	Link E	3.4%	Low	-0.8%	2.6%
6th	Link F	2.9%	Low	-0.6%	2.3%
7th	Link G	2.5%	High	0.5%	3.0%
8th	Link H	1.5%	Low	-0.3%	1.2%
9th	Link I	1.3%	High	0.2%	1.5%
10th	Link J	0.5%	Low	-0.1%	0.4%

This information allows the search engine to improve on its Initial Ranking by ordering the results according to the corrected probabilities. This Enhanced Ranking is shown in the table on the following page.

Table Comparing Initial Ranking of Links with Enhanced Ranking Based on Relevance Score of Links						
Initial Ranking of Links				**Enhanced Ranking of Links**		
Link Position	Link	Corrected Probability of Being Clicked		Link Position	Link	Corrected Probability of Being Clicked
1st	Link A	17.6%		1st	**Link B**	**20.9%**
2nd	Link B	20.9%		2nd	**Link A**	**17.6%**
3rd	Link C	5.1%		3rd	**Link D**	**6.8%**
4th	Link D	6.8%		4th	**Link C**	**5.1%**
5th	Link E	2.6%		5th	**Link G**	**3.0%**
6th	Link F	2.3%		6th	**Link E**	**2.6%**
7th	Link G	3.0%		7th	**Link F**	**2.3%**
8th	Link H	1.2%		8th	**Link I**	**1.5%**
9th	Link I	1.5%		9th	**Link H**	**1.2%**
10th	Link J	0.4%		10th	**Link J**	**0.4%**

Taking the expected probability of the user clicking on the link as an indication of the relevance of the link for the user, we can see that the Enhanced Ranking, based on the VortexDNA Relevance Scores of the links for the user, is a more accurate way of ordering the search or product recommendation results than the Initial Ranking.

More Information

Branton Kenton-Dau, Director
branton@vortexdna.com

Nick Gerristen, Director
nick@vortexdna.com

Raf Manji, Director
raf@vortexdna.com

Richard Waid, Technical Director
richard@vortexdna.com

