ALGO 119 - Module 3_2

Hello and welcome back to week 3 for this video lecture on content optimization and metrics.

Today we're going to talk about how publishers are working to optimize the distribution of their content on platforms and on their own websites and in their own apps.

I'm going to give you an overview of at least some of the techniques and approaches that are being used by news publishers to maximize the reach of their news content online.

The first thing I think it's important to understand here is that newsrooms are now using a lot of data to determine how content performs online. Tools with names like Chartbeat and NewsWhip and Parse.ly are used to collect that data and provide dashboards to reporters and editors to help them understand their audience better.

Now, once you're monitoring and measuring what the audience is interested in and clicking on and spending time with, that data can then be fed into data mining algorithms which can then be used to predict and optimize the performance of new content to maximize the reach of that content.

So let's look at a couple of the algorithmic data mining approaches that are now being used by news publishers to optimize content online. So the first technique I want to talk about is called A/B Headline Testing.

A/B Headline Testing is used to determine the best headline for a story with respect to some metric like say the click through rate for that story. In this image of the New York Times home page the headlines that are underlined in red are the ones that are being tested. Different people who load the page will see different versions of those headlines.

So the way it works is that journalists will write several alternative headlines for the story, and then the site will serve those different versions of the headlines to different people to see who clicks and how often they click on the different versions.

After some amount of traffic the testing algorithm will converge to indicate which version does best with respect to the metric. Again, that's usually the click through rate. So these techniques are now used at many different sites online and in fact sometimes the same optimization approach that's used for headlines is also used for images.

So at the Washington Post for instance, they also tried variations of thumbnail images as well as different blurbs for the stories. And in addition to that, they can even optimize for subsets of the audience to determine say the best headline and thumbnail combination to get clicks from some demographic group like say women or people in Canada.

Another technique that you might see is news organizations using widgets at the bottom of article pages which provide personalized article recommendations to users based on factors like the individual reading history, their interest in particular topics or what symbol or people have read.

These Widgets will recommend articles that are more likely to be clicked on by that user, and these widgets are really there to make the page more sticky, and to keep users circulating in recirculate into new content on the site.

At the Washington Post, the article recommendation algorithm increases clicks by about 9 percent in comparison to if you just randomly selected articles for that recommendation widget. Another content optimization approach has to do with figuring out the optimal time to publish a piece of content to a particular channel.

So there are some startups out there trying to do this for publishers like this one echo box that I'm showing you in the slide. There's another one out there called True Anthem that you can look up. They both use algorithms to data mine when your piece of content is most likely to get the most amount of attention online.

Now, one advantage to this type of approach is that you're using an algorithm to optimize against another platform's own curation algorithm. So remember in the last video lecture we talked about how frequently platforms like Google and Facebook changed their algorithms, if you use a technique like the one from Echo Box, this can help you be more responsive to changes in the platform algorithms because you're using your own algorithm to adapt to someone else's algorithm.

Now something to be wary of in using these types of automated optimizations is whether they're optimizing the appropriate metric for a particular piece of content or for your particular business strategy. So let's talk about metrics a little bit more.

The results of content optimization hinge on a key decision. What's the metric that's being optimized? What are you measuring about how people are engaging with the content on your site? Should your news site optimized for clicks or engage time or interaction with content or for other metrics like subscription, loyalty, overall satisfaction with the site or something else maybe? Are you really able to measure and optimize the things that you care about or will you work with the metrics that are convenient to optimize?

The data that's available from platforms is an important factor here, and it's another way that platforms can exercise their power over publishers. Metrics should also align with the business strategy of a news site.

If you're trying to optimize for subscribers, then maybe trying to measure the number of viral retweets on Twitter isn't going to matter as much as measuring something like the repeat visit rate and loyalty of users to your home page.
Metric should also align with content, and what I mean by that is that content might succeed according to different metrics. One piece of content might be meant to go viral and get a lot of clicks and another piece

of content maybe an in-depth investigation might be meant to appeal to potential new subscribers or appeal to loyal users who already subscribe.

In other words, you can't judge all content by the same yardstick or by the same metric. You have to think about what the content is supposed to do for your business or for your editorial strategy, and then measure the metric that helps you optimize for that.

So metrics need to be aligned for the editorial priorities of your news organization, and for its various types of content. That means that editors should strive to articulate the goal of a particular piece of content before they choose what to measure and what to optimize for that content.

Finally, I want to do a brief recap for today's video lecture. First, I'm giving you an idea of at least some of the ways that news organizations are optimizing content algorithmically including things like headline testing, use of recommendation widgets on article pages, and also optimizing the timing of publication of articles to different channels online.

Secondly, I hope I'm giving you some ideas for how to think about metrics with respect to content optimization algorithms. There's no one metric to rule them all, and you really need to think about optimizing the right metric for your content and for your business strategy.

I hope you'll also tune into my video interview with Tamar Charney from NPR this week. We'll be talking a lot more about algorithmic curation, and how to do that in an editorially responsible way. So I hope to see you there. Thanks.