SDS PODCAST
EPISODE 269: MAXIMIZING ADVERTISING EFFORTS WITH DATA

Show Notes: http://www.superdatascience.com/269
Kirill Eremenko: This is episode number 269 with Principal Data Scientist at ViralGains, Justin Fortier.

Kirill Eremenko: Welcome to the SuperDataScience Podcast. My name is Kirill Eremenko, Data Science Coach and Lifestyle Entrepreneur. And each week we bring you inspiring people and ideas to help you build your successful career in data science. Thanks for being here today, and now let’s make the complex simple.

Kirill Eremenko: This episode is brought to you by my very own book; Confident Data Skills. This is not your average data science book. This is a holistic view of data science with lots of practical applications. The whole five steps of the data science process are covered from asking the question to data preparation, to analysis, to visualization and presentation. Plus, you get career tips, ranging from how to approach interviews, get mentors and master soft skills in the workplace.

Kirill Eremenko: This book contains over 18 case studies of real world applications of data science. It covers off algorithms such as Random Forest, K-nearest neighbors, Naive Bayes, Logistic Regression, K-means clustering, Thompson Sampling and more. However, the best part is yet to come. The best part is that this book has absolutely zero code. So how can a data science book have zero code? Well, easy. We focus on the intuition behind the data science algorithms so you actually understand them, so you feel them through. And the practical applications, you get plenty of case studies, plenty of examples of them being applied.

Show Notes: http://www.superdatascience.com/269
Kirill Eremenko: And the code is something that you can pick up very easily once you understand how these things work. And the benefit of that is that you don't have to sit in front of a computer to read this book. You can read this book on a train, on a plane, on a park bench, in your bed before going to sleep. It's that simple, even though it covers very interesting and sometimes advanced topics at the same time. And check this out, I'm very proud to announce that we've dozens of five star reviews on Amazon and Good Reads. This book is even used UCSD, University of California San Diego, to teach one of their data science courses. So if you pick up Confident Data Skills, you'll be in good company.

Kirill Eremenko: So to sum up, if you're looking for an exciting and thought provoking book on data science, you can get your copy of Confident Data Skills today on Amazon. It's a purple book, it's hard to miss and once you get your copy on Amazon, make sure to head on over to www.confidentdataskills.com where you can redeem some additional bonuses and goodies just for buying the book. Make sure not to forget that step, it's absolutely free. It's included with your purchase of the book, but you do need to let us know that you bought it. So once again, the book is called Confident Data Skills and the website is confidentdataskills.com. Thanks for checking it out and I'm sure you'll enjoy.

Kirill Eremenko: Welcome back to the SuperDataScience podcast, ladies and gentlemen. Super excited to have you back here on the show. In today's episode we have a lovely guest, Justin Fortier, who's the Principal Data Scientist at ViralGains. That's gains with an n. He joined us for the
conversation today. What did we talk about? Well, ViralGains is a disruptive company in the space of ad tech, which is advertising technology. It's a very interesting space, because what they do is they show ads, such as video ads, for instance on YouTube or on other platforms, and they need to very quickly, within milliseconds decide which ad needs to be displayed.

Kirill Eremenko: For instance, they have multiple clients, and they pay for ad space. Then, based on who's watching the ad, they very quickly need to make the decision of what ad to display. Not only is there machine learning and algorithms involved, but also they have to happen very fast. Lightning fast, within a couple of milliseconds, because while the page is loading, you don't have much time to decide which ad. You have data science in a constrained space. You need to create algorithms and models, but you cannot... You don't have the luxury of time for those models to run within, like an hour or whatever else, they need to process very fast.

Kirill Eremenko: That's what we're going to be talking about here. We're talking about ad tech, performing insights, getting insights, making decisions within milliseconds with data science. You'll get some feature engineering ideas from today's podcast, which is always fun. You will learn about the business impact, why business impact is ultra-important for a data scientist to consider, and on the other hand, why user experience is an ultra-important factor for a data scientist to consider more and more in today's world.

Kirill Eremenko: You'll also learn about Justin's path, he's had a very interesting career from managing data scientists at

Show Notes: http://www.superdatascience.com/269
large organizations, to being the single data scientist at smaller startups. And he’ll comment on some interesting decisions he made throughout his career, such as picking an MBA over a PhD, so I found that quite interesting, and I think you will too. By the way, Justin and I met at the DataScienceGO 2018 Conference, so if you haven’t gotten your tickets for DSGO 2019 yet, make sure to check them out at datasciencego.com. On that note, without further ado, let’s bring our guest into the studio. I bring to you Justin Fortier, who’s the Principal Data Scientist at ViralGains.

Kirill Eremenko: Welcome to the SuperDataScience podcast, ladies and gentlemen, today I’ve got a super exciting guest joining us in from Boston, Justin Fortier. Justin, how are you going, my friend?

Justin Fortier: Doing great, Kirill, how are you?

Kirill Eremenko: Very good as well, thank you. Really cool to catch up, as I mentioned just before the podcast, when I saw your face I was like “Oh, you look familiar, where did we meet?” We actually met at DataScienceGO last year. Isn’t that crazy? How was the conference? How was your experience at DataScienceGO?

Justin Fortier: I loved it. I’ve been to a lot of data science conferences over the last five years or so, and that was definitely one of the best. It was super high energy, and exciting and motivating. Great speakers, and I was glad to meet you and Hadelin in person, because I had taken a bunch of your courses, and they had really helped

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
me in my career, so I wanted to shake your hand and thank you. That was great.

Kirill Eremenko: Yeah, thank you very much. It was really cool. I think you said we met at the yoga as well, in the morning. With [inaudible 00:07:17]

Justin Fortier: Yeah, I think there was a... I don't know if it was seven or eight a.m. out back, it was San Diego, so it was perfect weather.

Kirill Eremenko: Yeah.

Justin Fortier: We were doing some yoga, and I looked over. I was all sweaty, and I shook your hand. You were probably like "Who is this guy?" That was me.

Kirill Eremenko: That's very cool. How did you find the yoga, by the way? Jacqueline is a good friend of mine, and the point was for us... For her to help people energize before a long day of conferencing. Did you find that helped?

Justin Fortier: Yeah, it was great. In fact, I only did it one of the days, but I think, I wish I had done it more. It's a great way to start the day, and it wasn't like super hard-core exercise, but it was just enough to kind of get the blood pumping. She was great the whole conference, it wasn't just the yoga, right, she led us in a lot of cheering and singing and dancing.

Kirill Eremenko: Yeah.

Justin Fortier: She's one of a kind. You don't see her at most data conferences, that's for sure. That was a nice touch.

Kirill Eremenko: I'll tell you... Yeah, I'm in Bali now. Last year, that's where I met her, and we flew her all the way from Bali,
just for DataScienceGO, because of the energy she brings. Like, right here, she's... There's a place called Samadi Bali on the street I'm living on right now, and it's like a yoga studio. She does yoga there, and also she does these ecstatic dances, which is a way for you to release your inner self and dance to music. It's really cool, and when you see her doing one of these things... It's pretty much what she did on the conference, and yeah, the energy was incredible, so we had to bring her. It was good fun. I'm glad a lot of people have had great feedback about that component of the event.

Justin Fortier: You want to fly her from Bali to San Diego, and then have people have no idea who she was, right? She was very memorable.

Kirill Eremenko: Yeah, awesome, awesome. Speaking of courses, like you mentioned, you were excited to take... You had taken a few of our courses, of Hadelin and myself. You're doing your own course in Harvard in a couple weeks. Tell us a bit about that, that sounds super exciting.

Justin Fortier: Yeah, thanks, yeah. I live in Cambridge, five or ten minute drive from Harvard, and I work in downtown Boston. I met a guy named Ted [Kwartler 00:09:46] at a data science conference in Boston, I don't know, probably a couple years ago. He was working as a full-time data scientist at Liberty Mutual, an insurance company in downtown Boston. He was teaching a course for Harvard in their summer school, and he was looking for a guest speaker to come in and just
talk for an hour about how to use data science in the business world.

Justin Fortier: So, he had me come in representing advertising, and then he had another guy come in. It was actually using data science in a venture capital firm. I did the talk, and it was fun, and that was kind of the end of it. Then a couple of months later he contacted me and he said "Harvard’s looking to start a new course this coming summer." This would have been last fall, so they gave me plenty of notice. He said "Summer of '19 we're looking to start a course on data science in advertising and marketing specifically, and I think you'd be perfect. Are you interested?"

Justin Fortier: It's going to be a lot of work. It has been a lot of work. I've been working on it pretty much every Sunday for the last four months, and my wife's probably getting tired of it. But, yeah, it's a great opportunity. I couldn't say no to that. So, I'm going to be doing that this summer, and then they already have me doing one in the fall, and one next spring which will be more traditional. Once a week, that should be a lot easier.

Kirill Eremenko: Very cool. What are your students going to learn by the end of your three-week intensive course?

Justin Fortier: Yeah, it's kind of a combination of an overview of the advertising technology space. I think it's a field that kind of goes on behind the scenes, but a lot of people don't really know much about it. Then, I'm also going to mix in the specific ways that we use data science. I'm going to talk about some of the specific business problems that we have, and we're going to do some
coding in R and Python. I think it will be a good mix of theory and practical application as well.

Justin Fortier: It will be intense. Four nights a week, for three weeks. We're going to move quickly, but I think it will be great.

Kirill Eremenko: That's really cool. That's very exciting, and I'm sure it's going to go well. Actually, before this podcast, I had a look at your video called To Bid or Not To Bid. For everybody interested, it's available through Justin's LinkedIn, it's not public on YouTube, but if you go to Justin's LinkedIn and click on the link there, you can find it.

Kirill Eremenko: I was actually quite fascinated by the world of ad tech. I haven't thought I understand how these ads should work and the role machine learning plays, but I like the detail. The little detail you went into there. Specifically around the space of video ad tech. In a nutshell, can you just describe what is this industry all about? What is ad tech?

Justin Fortier: Yeah, sure, thank you Kirill. I have to be honest, I didn’t know a lot about ad tech either. The company I work for now is called ViralGains, and they specialize in video ad tech, as you mentioned. When they first contacted me, I heard about them from my head hunter about two years ago, and they were looking for a Principal Data Scientist. I just had a good feeling about it, I took the job and I've been learning every day since then. I still have a lot to learn, it's actually a very complex industry.

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
Justin Fortier: In a nutshell, basically, I think the easiest way to explain it is probably through a use case, right? When you open a browser and you go to your favorite website, whether it be on a desktop, or a laptop, or a tablet, you’ll see these ads popping up, right? Sometimes they're static ads, which we call banner ads, or display ads, and sometimes they're video ads. It could be anywhere from five-second video up to a two-minute video. We don't create the videos, but we work with companies like Toyota, Jet Blue, Amtrak. A lot of the big brands, a lot of the banks, all different industries.

Justin Fortier: Our job is to basically help them get eyeballs on that video by their target audience. They might come to us and say "We have this 15-second spot about a new Toyota Highlander truck. We'd like to show it to 10 thousand people like Kirill by the end of the month, go." Then I have to figure out from a data science perspective, do we want to bid on ad space when each individual person opens the browser based on their likelihood to watch the video? Because, we only get paid if you watch it for a certain amount of time.

Justin Fortier: The core machine learning behind it is the likelihood that you'll watch it for a long enough time, but there's also a lot of other things. You can share it with your friends through social media. We do other things, like if you click and go to the website and then purchase something, we can measure the ROI on that. It's all about getting the right... We don't create the video, but it's all about optimizing the process so the right video
gets seen by the right person at the right time on the right device, so they’re most likely to act on it.

Kirill Eremenko: Hmm. Very interesting. You’re mostly working, I’m assuming with classification models. Is that right?

Justin Fortier: Exactly. It’s been mostly classification up to this point. We’re actually doing something now, which is really cool. One of the things we do is we have different engagement experiences. We don’t just show you the video, we’ll do other things after the video. One of them is a survey, for example, right? We’ll say "Now that you’ve watched this video, how likely are you to buy a Toyota in the next six months?" For example. It’s a one to five star likelihood survey question.

Justin Fortier: We’re able to compile the results and send them back to our client, which would be Toyota in this example, and tell them for future targeting purposes, these are the people that are interested. They tend to be in the geography, or this age group. All different ways we can look at data. One of the issues with surveys, particularly if you don’t incentivize people, is the response rates are extremely low.

Justin Fortier: One of the things we’ve been working on the last couple of months is... Let’s say only one percent of people respond to a survey, how great would it be if we could build an audience, a look-alike model to infer what the other 99% of people that were exposed to this survey would have answered if they answered the survey. Based on what we know about the one percent that did answer.
Justin Fortier: That's something that's really exciting. No one else in the industry is doing it, and that'll probably be a future presentation. Talk about how we did that. It's pretty cool stuff.

Kirill Eremenko: That's awesome. It's kind of like another machine learning model on the side, going off on a tangent. You got your main one doing the predictive part, and then on the side you're using machine learning to describe the results of your machine learning project.

Justin Fortier: Yeah. It's sort of like a multi-class classification problem, right? Instead of your normal classification problem where it's binary, your target variable is yes or no, one or zero; in this case we're trying to predict will they answer one star, two star, three star, four star, five star. Right?

Kirill Eremenko: Yeah.

Justin Fortier: It's a multi-class classification problem.

Kirill Eremenko: Gotcha.

Justin Fortier: It's more complex, but I think it's more interesting too.

Kirill Eremenko: That's really very interesting, what you said about the whole you guys don't get paid unless a person watches a certain amount of video. In your presentation, I think that was one of the key moments for me. I think a lot of businesses out there are tired of advertisers and marketers. They just take your money and... Again, I'm speaking as a business owner. They just take your money and are like, all right, we promise you these things, you'll be exposed here and there and so on. But, they never commit to certain KPIs, certain
actual measurable ways. It takes months, sometimes
even years... Not as long as years, but up to a year to
figure out this is not actually bringing us any value.

Kirill Eremenko: It's giving your brand exposure and so on, but people
are not really taking action from these specific
marketers. The input they're doing for the company.
What I liked in your video, like what you mentioned
just now, that you guys don't get paid unless a certain
amount of video's are watched. Then, I think it was
either you, or your co-presenter, Michael that said you
guys took it even further. Now, you're starting to look
at, apart from minutes watched and clicks, you're
starting to look at ways for different companies...

Kirill Eremenko: Like for instance, for an auto manufacturer, with the
right data points you can measure whether or not that
person actually scheduled a test drive with a car.

Justin Fortier: Right.

Kirill Eremenko: Or, if somebody walks into a store, if you're doing it for
a retail store. Somebody walks into a store, using with
their consent of course, using their information from
their mobile phone, you can determine if they came
into the store based on watching an ad or not. You
can't determine that, but on the client side they can do
that. Therefore, they can create this feedback loop. Tell
us a bit more about that. Is that where the space of ad
tech is moving to?

Justin Fortier: Yeah, I think it is. I think marketers, publishers, you
know there's a whole complex environment around
this stuff. I think they're moving beyond just trying to
get eyeballs, or even just trying to get clicks.
Ultimately, they want a business outcome, right? They're spending so much money to work with us, and they want a return on that investment, you know?

**Justin Fortier:** That could be defined many ways, and we get different clients from different industries that define it dozens of different ways. For one, it might be... Ultimately, it comes down to placing a conversion pixel on the website. It might be... If it's a non-profit, it might be after Justin watches this video, we want him to fill out this form for more information and click submit. We worked with University of Phoenix, right, their whole thing was don't just watch the video, actually click submit after filling in some basic demographic information, contact information saying that you want to know more about our school. Then we'll follow up with you. For them it was a lead generation thing.

**Justin Fortier:** As you mentioned, for auto, it could be totally different. It could be go to the website and schedule a test drive. Only once they do that does it fire the conversion pixel and we can say that as long as it happened within a certain number of days after they watched the video that we showed them, then it's attributable back tot hat video. Right? You know, it can't be something that happens, like six months later, it has to be within a week or two. It's however the client defines if, and that's why we have to be flexible in terms of how we build models around this stuff.

**Kirill Eremenko:** Gotcha.
Justin Fortier: One thing I wanted to mention if I could, quickly, was not only do we not get paid unless they watch for a certain amount of time, but in terms of... I mean, you own a business, you understand cash flow, right? Cash in minus cash out is your profit. Well, we have to pay up front, so say you open a browser on your favorite website, there's what's called real-time bidding. There's an open real-time auction, and we're what's called the demand-side provider, but there's also supply-side providers that supply the inventory.

Justin Fortier: The inventory is just an ad space. Think of it as you open your favorite website, a notification is sent to us, and all the other DSPs that we're competing with. It says "We have this guy, Kirill, this is where he's located, this is what device he's on, what operating system he's using. This is the website he's on. Some basic stuff we know about him. Are you interested in buying this ad space?" If so, go ahead and bid, and it's a blind auction. We could bid and lose, or we could bid and win.

Justin Fortier: Another optimization game is we have to figure out what's the right amount to bid. Because, we don't want to bid too much, where we could have won it for 10, and we spent 20. But, we also don't want to bid nine and just barely lose out, right? Because it's a blind auction, that's a whole other purpose for data science, to figure out the market value of that ad, specifically for you.

Justin Fortier: Then, if we decide we're going to bid on that ad and we actually win it, then we have to decide which video do we show you first, right? You might qualify for three of
our target clients. Jet Blue, Toyota, and Amtrak, let’s say. Which one do we want to show you first? Well, the only way to know that is to calculate the likelihood that you’ll watch each one of them for long enough for us to get paid. So, it’s three separate models just for you.

Kirill Eremenko: Mm-hmm (affirmative). Wow. All of that has to happen in less than a second.

Justin Fortier: Yeah, it’s like 10 milliseconds, yeah. It all has to be automated and at scale. That’s why I build the models initially, and then I work closely with the software engineering team to deploy them and production-ize them, and make them happen fast enough.

Kirill Eremenko: That’s crazy. We’ll get back to them in a second. I just want to confirm that part where you said not only do you not get paid, but you actually have to pay. What does that mean?

Justin Fortier: Exactly. Yeah, basically, we’re bidding on that ad space. It doesn’t cost anything to bid, but if we win, then we have to pay for it, right? We’re paying up front, then if you don’t watch it for long enough, we don’t get any money, so that’s basically lost profit on you, right?

Kirill Eremenko: Oh, so hold on. You pay... Doesn’t the client pay for the ad space?

Justin Fortier: No. They don’t. That’s on us.

Kirill Eremenko: Do you take on all the risks? I want to work with ViralGains now.
Justin Fortier: Yeah, yeah.
Kirill Eremenko: Send me the pricing.
Justin Fortier: No, we take on the risk. Not only that, but if... There's sort of... think of it as the micro-goal is you individually. Are you going to watch it long enough so that we get paid?
Kirill Eremenko: Yeah.
Justin Fortier: There's also macro goals which are, remember, Toyota said we want to have 10 thousand people like Kirill watch this in the next month.
Kirill Eremenko: Yeah.
Justin Fortier: We can't just say "Well, we lost this one. We had a bad week, but we'll make it up next week." You know? They check us every hour of every day to make sure that we're on pace for that monthly goal. It's pretty intense and pretty complex. Way more than I ever imagined.
Kirill Eremenko: Wow, wow. Very interesting. Tell us a bit more about this whole speed situation, because in your presentation I was fascinated about this. That unlike other industries that are using data science, you have this additional constriction of time. 10 milliseconds is not a lot of time at all, so one of the interesting things that you guys mentioned in your talk was that you can't afford to use deep learning.
Justin Fortier: Right.

Show Notes: http://www.superdatascience.com/269
Kirill Eremenko: As much as you'd like to, deep learning is not an option for you because, just, you don't have the time for those neural networks to calculate.

Justin Fortier: Yeah.

Kirill Eremenko: Tell us a bit about that. How is it working in a constricted type of data science machine learning space?

Justin Fortier: Yeah, I think it's very true. That's why I work closely with the engineering team that's deploying the models. Sometimes they'll come back to me and say "There's a latency issue," which basically means it's too slow. It's taking too long, not so much to train the model, but to score. They train it nightly, but then they score people in real time. So, when you open a browser, they're scoring you right away, in fractions of a second. To your point, we might not be able to do a neural network with 17 hidden layers or something, we might be stuck with more like...

Justin Fortier: You know the logistics of it. I mean, I've used on some of the models, I've used XG Boost. We've been able to do that, so it's not like we're stuck with a very basic simple regression, there are still a lot of things we could use. Yeah, it has to work quickly, so that's another consideration we have to think about.

Kirill Eremenko: I guess it's also on you to be smart about not only what you use but how you use. You know, how you structure your code, how you make it fast. What are some of the tips you have for somebody in a similar space that has a similar problem? They have this time restriction, time constraint. What are some of the tips
or hacks that you can share for getting the results fast?

Justin Fortier: Yeah, sure. I think Occam’s razor, right, keep it simple is always a good one. When I’m thinking about feature engineering, I might start with 100 variables, but in the end, the final model might have 10 or 15 at the most. It’s really trying to figure out which are the features that are most predictive of your target variable, and only sticking with those. Not including other variables that will just, basically, make it overly complex, increase your chances for over-fitting and then slow down the model. Those are three negative things that can come from a model being too complex.

Kirill Eremenko: Mm-hmm (affirmative). Feature engineering, getting the right amount. I guess that’s probably one of the biggest ones. Is there anything else that really reduces the load, or the latency?

Justin Fortier: I think that’s probably the big one. Again, like you said earlier, sticking to certain models. Avoiding overly complex models. When I say model, I’m talking about the algorithm, right? Maybe you’re not able to use all the most complex algorithms, but you’re still able to use many different algorithms, and then the model itself, if there are ten features that give you 99% of the performance that the extra five features would have given you in half the time, then it’s a cost-benefit analysis. Maybe you don’t need the extra five variables, you can do the same thing with the 10 that you have.

Justin Fortier: It’s not... One of my favorite quotes is "Perfect is the enemy of good," right?
Kirill Eremenko: Mm-hmm (affirmative).

Justin Fortier: I myself, as I'm sure a lot of data people are, we're sort of type A perfectionists. One of the things I've learned throughout my career is the best model in the world doesn't give you an ounce of business value, unless it's in production. Even then, sometimes it takes months to realize the business value, depending on the model. Spending that extra week to get an extra 0.01 on your area under the curve, or an extra one percent on your precision might not be worth it. It might be better to just get something that's really good and get it out there, and start achieving value from it.

Kirill Eremenko: Gotcha. I totally agree. Tell us a bit about refreshing models, how often do you re-write, retrain your models?

Justin Fortier: Yeah, we retrain them... Are you talking about when we actually run them... we retrain them on a nightly basis. Based on all of the impressions that came in, in the last 24 hours. When I say impressions, those are people that are opening browsers, right? We call them impressions. We retrain on a nightly basis. We talked about that.

Justin Fortier: That's one of the interesting decisions that we had to make. Is that enough? Should we do it hourly? Should we do it real-time? Should we do it once a week? The short answer is test out different things and see how often it actually changes. You don't need overkill, right? If once a night is fine, and it's not going to change that much every hour, then that's probably sufficient.
Kirill Eremenko: That's very interesting. We were actually talking about training a model as in, you have this data from during the day, who clicked, who didn't. You want to go and update your coefficients and things like that inside your algorithm, so the next day it's applied.

Justin Fortier: Right.

Kirill Eremenko: Most companies, or people who I've spoken to, or know how their models work, most companies don't even do it that often, you know? Many banks will do it once a month, or once every two months, retrain a model.

Justin Fortier: Yeah.

Kirill Eremenko: So, does the dynamic really change on a daily basis? What is that assertion?

Justin Fortier: It does, and I think... We looked at that, and we ultimately decided that it changes enough day-to-day that it's worth retraining the models, if it wasn't an extra load in terms of engineering. Right? It really didn't make any difference to them, or to our systems, so we ultimately decided that once a day for this was good. It changed enough on a daily basis.

Justin Fortier: You have to remember, we're talking about crazy volume here. We literally look at billions of impressions every minute.

Kirill Eremenko: Wow.

Justin Fortier: Right? It actually... It's not a small retail mom and pop shop or something where we might get 100 transactions a day or something, right? The type of people that buy frozen peas probably didn't change
much from yesterday to today, right? But, when you're talking about seeing billions of people a day, it actually changes pretty frequently.

Kirill Eremenko: Interesting. Do you think... What do you think that change is associated with? First thing that pops to mind to me is, what if it's just day of the week? Maybe on Tuesdays people are less excited to click on stuff. But then, in that case you would have a lag.

Justin Fortier: Yeah, we've actually-

Kirill Eremenko: No?

Justin Fortier: Yeah. We've actually looked at day of the week as a model feature too. Right?

Kirill Eremenko: Okay.

Justin Fortier: That becomes a categorical variable with seven values. Or, hour of the day becomes a categorical with 24 values. Then you have to convert it to local time, right? Because, you don't want to compare what you're doing at seven a.m. in Bali and I'm doing at seven a.m. in Boston, right? That's two totally different times. Yeah, there are a lot of things it could be, that's why we...

Justin Fortier: When I first start building a model, I look at basically every potential variable that I can think of and the business can think of. I meet with the executives and "What could possible help us predict this target variable?" We kind of whittle it done from there.

Kirill Eremenko: Mm-hmm (affirmative), mm-hmm, okay. What would you say are the factors that make this space so dynamic? Why is it changing on a daily basis? Is there
something in the audience, is there something in the product, is there something in-

Justin Fortier: Yeah.

Kirill Eremenko: Where? I’d love to know, or get a feeling for what can make people’s opinions... At the end of the day, people on one day they’re more likely to click if you have these coefficients this is how you serve the ads.

Kirill Eremenko: Then, the next day something has changed about them, like their moods, their attitude towards a product, that now you need to serve them a different ad, or you need to serve them a different product overall. I’m just curious, what are your thoughts on that?

Justin Fortier: Yeah, sure. Yeah, that’s what makes it so interesting right. On a macro level, some of the things that are changing, and this is actually some of the stuff that I’m going to talk about in my upcoming course. First of all, with the whole Facebook data privacy concerns that are going on, people are very leery to do anything, because they're kind of worried that you're going to find out too much about me. There's personally identifiable information, and I don't want to answer any questions. What happens if I watch this video?

Justin Fortier: There's something called ad choice, right? At the bottom right, which you can opt out. There are concerns like that, where we're up against it in the sense that I think everybody is more cautious now than they were even a year ago, in terms of what they do online, and who's going to see it, and how they're going to be tracked and targeted based on that. But,
that said, some of the things I think are the type of the video itself. What is in the video? Is it a persuasive video? Is it an instructional video? Is it meant to be funny? Are there famous celebrities in the video? Is there music? What kind of colors are used?

Justin Fortier: There are a lot of things you can look at in terms of the video itself. If you were to ask me, most people, let's be honest, they don't like those ads popping up, right?

Kirill Eremenko: Yeah.

Justin Fortier: You know, our CEO's mission is to show people ads that they don't think suck, or something like that. I'm paraphrasing, but we want to suck less than other people, right? It's a numbers game, right? It's a percentage game, so what would you guess, if I were to ask you what percent of people do you think watch a video for... Let's say it's a 30 second video. What percentage of people would you say watch it in its entirety?

Kirill Eremenko: In its entirety? I would say 10, 12 percent.

Justin Fortier: Last year, we were able to get that up to 75%.

Kirill Eremenko: No way.

Justin Fortier: Yeah. That was a pretty good result. Like you said, anecdotally, when you talk to 10 friends, they're all going to say "No way, I don't watch those things," right?

Kirill Eremenko: Yeah.

Justin Fortier: But, there's something that you would watch. Is it the video, is it the mood you're in? Is it demographically?

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
Certain types of people like certain types of videos. That's what makes it so interesting. Honestly, I wish we had more data than we do sometimes. There's a lot of data about the person and the personality, and things like that. You know, we have some, but if we had more, I think we could do even better. But, I think based on what we have, we're very happy with the results we get.

Kirill Eremenko: That's really cool, 75% is incredible. You just made me think about my past couple of weeks, what kind of ads I've been seeing, and indeed you're right, there are certain ads I will watch to the end. For instance, if I happen to log on to Instagram, and there they often put in ads.

Kirill Eremenko: Recently I've been thinking about buying a backpack, so I've been searching for backpacks. My backpack is 4 years old, and it's time to update, so I've been searching for an advanced backpack. You know, waterproof with certain security features and so on. Now I'm getting these ads for backpacks, but that's actually very handy for me, because I don't have to search for it myself.

Kirill Eremenko: I'm getting these Kickstarters that I might be interested in, or these really cool backpacks that you can already buy and get shipped. That's a really cool way for me to save time, because it's all summarized in that video. I watch 30 seconds, if I'm interested in more, I just go and click on it, so I totally get what you mean. Sometimes I don't mind these ads, if they are serving me, and they're useful. It's really cool-
Justin Fortier: You know, I think the word that comes to mind is relevance. If it's relevant to you and your specific situation, you're more likely to engage with it, right?

Kirill Eremenko: Mm-hmm (affirmative).

Justin Fortier: If it's not, then you're not, right? It's interesting, one of the things I was just thinking about when you were saying that Kirill, was that's part of our decision too. That's a good thing that we have several clients at the same time. You might qualify for the target audience of five of our different clients, right?

Kirill Eremenko: Mm-hmm (affirmative).

Justin Fortier: Well, that's five different videos that we could show you. You know, one of them you're going to be more likely to watch than the other four. Right? So, we have to figure out which one is that and show you that one first.

Kirill Eremenko: Yeah, yeah, that's totally-

Justin Fortier: It could be a backpack in your case, or golf clubs, whatever it is. But, based on what we know about you and people like you, we have to maximize the likelihood that you are going to watch the ad.

Kirill Eremenko: That all kind of ties in, in the end... Again, something that you mentioned in your presentation that really resonated with me is user experience. Right? We are beyond the days where, like 2000's, or probably early 2010's, like 2012, 14, and so on, where people were just being bombarded with these ads where you couldn't care less. You just click close and so on. It still happens on YouTube quite a bit, where there's an
ad, "I just want to watch the video, leave me alone" type of thing.

Kirill Eremenko: But, now, we're moving to a space, or a time when user experience is at the top of the priority list. That we want the people who watch this to have a great time. We don't even just want to throw those ads out there just to meet our KPIs, or getting 10 thousand impressions. We want to actually make those impressions count. We want people... If a person's not interested in buying a car, why would we show him this video, right?

Justin Fortier: Right.

Kirill Eremenko: We want to maximize as much as we can the information that we have, the algorithms that we're using, in order to create the ultimate user experience. At the end of the day, imagine... I would love to see a world where not only the advertisers are happy because their ads are getting out there, but users predominantly... Like, 75% would be amazing. If 75% of people are happy that they're getting the ads that they're getting. They're serving them. How cool would that be?

Justin Fortier: Yeah, exactly. That's what we're trying to do. Another macro trend that I just thought of, besides the privacy concerns, is everything moving to mobile, right? It could be a very different experience too, some people might be more likely to watch a full ad if it's on their desktop, and it's a bigger screen. Other people might be more likely to watch it if they're on a mobile device, because they have an hour commute every morning on
the train. They have a tablet, or they have a phone, and once they get to the office, on their desktop, they don't have any time for that stuff. But, on the train they might be more willing to watch it.

Justin Fortier: That's another thing, the movement toward mobile, and everybody spending, whatever the crazy average is. Like, four or five hours online every day on their mobile devices. How do we have to adjust our business, based on that trend? There's a lot of interesting... It's always changing, it's never the same. That's why I like it.

Kirill Eremenko: Yeah, you guys serve ads across all platforms, including mobile?

Justin Fortier: We do, yeah.

Kirill Eremenko: Awesome, awesome. Another thing that was interesting that you mentioned before was that you have parameters about the video. Tell us a bit more about that. You not only have parameters or features about the person or the time of day, what type of device they're watching on, but also you use... Whether it's a car video, whether it's a cheap product, an expensive product, I'd be curious to know which features you use from inside the video, to describe the video, in your model.

Kirill Eremenko: It's very interesting when you have both sides, just selling a single individual product to a person and therefore you can only use the features about them. Like, let's say, end-user features. But, now you have multiple products that you can feature to the user. As you said, you want to show the one that will maximize

**Show Notes:** [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
the outcome, so now you can have features from the start. The starting end features, and the final features.

Justin Fortier: Yeah.

Kirill Eremenko: That sounds really cool. Not many data science problems have both sides involved in the feature engineering. Tell us a bit about the features.

Justin Fortier: Yeah. I think it doesn't exist is the short answer, right? It's not like Toyota will send us a 30-second video and they'll say these are 20 attributes about the video, and we just add them to our database and we have a nice, handy attributes table that we join to our other tables about the user and the environment, and those kinds of things.

Justin Fortier: We kind of have to build it ourselves. Some of it is proprietary, I can't give specifics, but yeah, the idea is we need to classify each video ad that we receive from a client according to several attributes. To your point, it's one thing to say that people over the age of 70 never watch videos, so we're just not going to show it to them. Well, there are two problems with that, right? One is our client might come to us and say "We want you to show this to 10 thousand people over the age of 70." You know?

Kirill Eremenko: Mm-hmm (affirmative).

Justin Fortier: We can't come back and say... We could say "We don't think that's a great idea." Or, "This type of video tends to do better with a younger demographic," that kind of thing. Ultimately, they're designing ads. They have experts that design ads. They don't necessarily want
our feedback on how to build a better ad, so what we do have to do is optimize with what we're given. We can say "This video has these attributes, and we've noticed historically that this video doesn't do as well with this type of demographic."

**Justin Fortier:** But, like I said, you can either target a different demographic, or you can change the video slightly to include softer music, or a different spokesperson, or the demographic loves pets. You know, you should have a pet in your video. As you can imagine, there's no blueprint, right? I mean, I looked for days online for anything like that, and it just doesn't exist. In a way that's exciting, because that means that we can build something that nobody's done before.

**Justin Fortier:** It's a lot of hard work, but when you find those two or three features from the video that help the model performance incrementally over what you already have, that's sort of the eureka moment. It might be that it's not just that these things help overall, but these particular video features help with a particular demographic that you're having a hard time reaching. Even if they don't work with other demographics, they work well with this demographic. That's why I say it's an optimization problem.

**Kirill Eremenko:** Gotcha. Just so I understand a bit better, the way it would work I, for instance in one of the... You have five, four, three products that you can potentially choose from. In one of the ads, let's say you have a pet. Then, not across every single model. You wouldn't put the categorical Boolean variable has pet or not, you wouldn't put that... You wouldn't include that
information as a categorical variable in every single model, you would just include it in the model that actually has the pet. Then you would see if that... what kind of demographic that helps with, right?

Kirill Eremenko: For instance, for some sort of demographic, let’s say age groups between 20-30, having a pet doesn't matter. Or, makes them less likely to see the video. But then other groups, like 30 and above, or under 20, having a pet in the video will make the model spit out a higher percentage.

Justin Fortier: Right.

Kirill Eremenko: Is that how it would work?

Justin Fortier: Yeah. There are a couple things. One, I think you could have different models. You could have a one-size-fits-all model that you build on all of your historical data, across all videos that you've shown historically. All video attributes, and all types of users. Overall, it will tell you these are the five things that really matter the most. In terms of somebody’s likelihood to watch your video, or to click, or to submit a form, or to purchase something online, or to schedule a test drive. What ever the business outcome that you’re trying to encourage is.

Justin Fortier: But, if you find, to your point, through exploratory data analysis that... If you picture a heat map, where you have down the first column, five different buckets for age. Under 18, 19-25 et cetera, et cetera. Then, across the top, you have maybe 20 different video attributes, right? Then, in the cells, you have the average number of seconds that somebody watched a
video. Or, the likelihood that they watch it for 30 or more seconds, something like that.

Kirill Eremenko: Mm-hmm (affirmative).

Justin Fortier: You know, then you can look for the red areas and say, "Okay, look, these three attributes do really well with the younger demographic. The older demographic, they don't move the needle at all." Then, from as business perspective, you can say "Do we want to include those in the overall model, or do we want to build a separate model by demographic group where they have different attributes for each?" The more specific you get, the better your model performance is going to be.

Justin Fortier: The issue, of course is-

Kirill Eremenko: The time.

Justin Fortier: You don't want to go too crazy, because then, in that 10 milliseconds where we have to decide do we want to bid on Kirill, how much do we want to bid on him, and which of these five videos do we show him... That's already five models, right?

Kirill Eremenko: Yeah, yeah.

Justin Fortier: We don't want to exponentially be increasing that, because then it takes too long.

Kirill Eremenko: Gotcha. Wow. Very exciting. Well, Justin, huge thank you for all the description of how this whole ad space works. At this stage, I think, let's switch topic a little bit, because you have a very interesting background, and I would... We can keep talking about ad tech
performing for ages, but I would love to talk about your background as well, your journey in data science. Because, I think a lot of our listeners can get tons of value out of that as well.

Kirill Eremenko: Let's start with the whole notion of you having... I think this is very interesting. That you actually don't have a physics or hard science background, as you say, but you do have an MBA. What is it like to get into data science with an MBA? Although that's probably not the right question, because you've been in data science for 20 plus years. Even, before the term was coined. Walk us through this. How did you get into data science? How did the MBA fit into all of that? What's your career been like?

Justin Fortier: Sure. Sure. Yeah, so I'm a little different than some people in the sense that I know a lot of data scientists come from PhDs in physics, or statistics, or something like that. I always have had an acumen for numbers, I did major in math undergrad. It's no like I came from an archeology major or something, you know?

Kirill Eremenko: Mm-hmm (affirmative).

Justin Fortier: I've always worked with numbers, long before it was called data science. I was always in analytics and things like that. At some point, like you said, midway through my career, I kind of had to decide, to get to the next level, to be the director at the VP level, to be managing teams of data scientists, would it be more beneficial for me to go back to school and get a Masters in stats and then a PhD, which would take, like six years? Or, to get an MBA? I just decided that
at the time, based on my situation, my family and everything else, that the MBA would have a higher return on investment for me. That's why I did it.

Justin Fortier: I think the MBA has been great. I went to Babson, which is a pretty technical school near Boston. There's a lot of entrepreneurial focus, so it really helps you in terms of thinking about the business first. All your decisions in data science are don't start with the model and say "I heard a support vector machine is this really cool model, I've got to figure out how I'm goin to use it." You know, that's kind of a backwards approach, in my opinion. The problem is, we're losing 10% of our customers every month, we need a model that's going to help us improve retention by 20%, or something like that. Now, which models can help us do that?

Justin Fortier: It's starting with the business problem, and then determining which model you want to use. Rather than the other way around. I think I got that from my MBA, just kind of thinking about it that way. Ultimately, no matter what you do in data science, with a few rare exceptions, you're going to be working for somebody else. They're going to have business goals, and they don't care what your area under the curve is, or they don't care what a confusion matrix is or anything else. They just want to be able to say "What was the business impact based on your model? How quickly did you get us to that point?"

Justin Fortier: I think the MBA has been great for me in terms of that. Also, just in terms of managing, presenting, communicating with executives. All those things I

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
think you get from an MBA that you might not get from a PhD in a hard science. That's just my opinion, you know?

Kirill Eremenko: Yeah. I know, gotcha. You mentioned in your current role, you do need to communicate with your clients and executives as well, to make sure you are on the right track in terms of that business impact with your ad. [inaudible 00:53:25] help.

Justin Fortier: Yeah, and it's all... If they ask you a question it's good to understand why you're doing this, and you can make some good suggestions back to them, and feel like you're contributing on a strategic level as well as on a technical level.

Kirill Eremenko: Gotcha. You've also worked in a wide diversity of industries and company sizes. I'd be curious, you worked at Staples for example.

Justin Fortier: Mm-hmm (affirmative).

Kirill Eremenko: That's a massive company with thousands of employees. Now, you work at ViralGains, which is more of a startup, smaller size.

Justin Fortier: Mm-hmm (affirmative).

Kirill Eremenko: What is the feel, in terms of doing data science at companies of different sizes?

Justin Fortier: Yeah, it's interesting. Staples... I think there are pros and cons to both, right? I also think that... Honestly, it wasn't like it happened by design, right? I just sort of... You do this stuff long enough, you end up working for different industries, different sizes of companies. I
think one of the reasons that I was interested in the startup this time, honestly this is the first true startup that I've worked for... It's just that things move quickly, right?

Kirill Eremenko: Yeah.

Justin Fortier: I think, sometimes when you're with a big company, I've had the experience where you build this really accurate model, but by the time it gets deployed, and by the time the business realizes any value from it, it's six months, eight months later, and by then so much has changed in the environment that your business is in that the model is not even relevant anymore. Right?

Kirill Eremenko: Mm-hmm (affirmative).

Justin Fortier: You know, that gets frustrating. I love the culture where I am now. In a start up it's sort of fail quickly, and there are a lot of experimentation, you know, come up with a hypothesis, test it, build a model, get it out there. We can always improve it later. Just move quickly, and I like that. You know, I like that.

Justin Fortier: I think one of the advantages of a bigger company, particularly in data science, is you can learn from other people. I'm sort of the guy. There has been several jobs where I've been the first guy, then I built a team, and that's great, and I enjoy that. I certainly learn from the people that work on my team, but it's not the same as working at an Etna, or a Wayfair, a couple examples of big companies in Boston that have hundreds of data scientists. Where every day you're sharing code with 10 brilliant people, and you're learning from them.

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
Justin Fortier: You know, that’s something that you do miss in a startup. I think there are pros and cons to both, but I’m pretty happy where I am right now.

Kirill Eremenko: Awesome, awesome. Do you think you're going to grow your data science team, because I'm 100% sure there are data scientists in Boston listening to this who would be super eager to work side by side with you. It sounds like you could learn a lot. Heck, I would work in your team just to learn all these things you're talking about. Any plans in growing the team at ViralGain?

Justin Fortier: Yeah, I think so. Thank you for the kind words. I think... You know, we are a startup, and the budget's tight, so it's not like we're hiring 10 this year or something, like some companies might. You know, I wouldn't be surprised if in the next six months or so we start looking for somebody.

Justin Fortier: I would also just throw out there that I'm perfectly happy to jump on a call with somebody, or have coffee with somebody if they're local. Just talk about their career and any ways I can help them, even if it's not necessarily... Even if we're not hiring right now, I know a lot of people in data science at all different companies around the area, so I can maybe get them in contact with somebody that is hiring.

Kirill Eremenko: Fantastic. Speaking of that, what are some of the best ways to get in touch with you, Justin?

Justin Fortier: Yeah, I'm on LinkedIn. Then, I guess, e-mail, if they wanted to, they could reach out to me on e-mail. It's

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
just justinfortier111@gmail.com. J-U-S-T-I-N-F-O-R-T-I-E-R-1-1-1@gmail.com.

Kirill Eremenko:  Gotcha.

Justin Fortier:  I'll hire anybody except you, Kirill. Just joking. You're unrecruitable anyway, that's what it says right on your profile. You're unrecruitable. I won't even try.

Kirill Eremenko:  Yeah, you know, I would love to, but... As you say, the time thing, right? You're feeling the same thing now. You've got a job, you've got a family. How many kids do you have?

Justin Fortier:  I have seven-year-old twins, and a 24-year-old musician down in Brooklyn. Trying to make it as a professional musician. We have two dogs, an old dog, and a new puppy that eats everything in sight. There's never a dull moment in the Fortier household, for sure.

Kirill Eremenko:  Yeah. On top of all of that, you're about to tach a course at Harvard. Heck, where does the time for this come from?

Justin Fortier:  I know. I know, but how do you say no to that, right?

Kirill Eremenko:  Yeah.

Justin Fortier:  I like to challenge myself, and sometimes you've got to go outside of your comfort zone a little bit. Do my best.

Kirill Eremenko:  By the way, where can people find this course? Like, if somebody wants to sign up for it, if it's not too late, where can people find it?

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
Justin Fortier: Yeah, so it's in what they call the Summer School. They have three-week intensive programs, and they also have seven-week. This is the first of the three-week programs. I don't know the exact website, but if you just Goggle Harvard Summer School, you'll see it. The name of the course is... I think it's called advertising and real-time technology. Something like that. If you just Google Harvard Summer School advertising, you'll see it. And, you'll see my name as the instructor.

Justin Fortier: I'd love to have anybody that's interested sign up. We still have quite a few seats open. It'd be great.

Kirill Eremenko: Yeah, guys hurry up. By the time this podcast is released there might not be quite a few seats open.

Kirill Eremenko: Awesome, awesome. Well, Justin, thanks so much for coming on the show, I just have one question left for you: What's a book you can recommend to our listeners that might help them to make careers in their lives?

Justin Fortier: Yeah, sure. I really like... There are so many good books about data science, but I'm all about the practical application, the business application of it. So, for me, one of the books that really... I read it three or four years ago and it really got me interested in this stuff. There were a couple. One was called Predictive Analytics, by Eric Siegel, who runs one of the local conferences here in... Predictive Analytics World, I think it's called here in Boston. That was a good book in terms of data science 101 without being super technical.
Justin Fortier: Another one I like is Data Science for Business, it's called. The authors are Foster Provost and Tom Fawcett. That's really good. It gets into a dozen or so specific business use cases where you'll need data science, and it also gets into more of the code and the technical approaches to tackling those problems. Those are two really good ones.

Kirill Eremenko: Thank you, thank you. So, Predictive Analytics, and Data Science for Business. Yeah, on that note, Justin, thanks so much for coming onto the show, sharing insights. I really hope I'll see you at DataScienceGO again this year. Let's stay in touch, man.

Justin Fortier: Sounds good. All right, thanks Kirill.

Kirill Eremenko: Thank you everybody for being part of today's episode. Super pumped about the conversation we had. I hope you enjoyed this [exscourse 01:01:38] into the world of ad tech, what an incredible world. You got to do data science, and you got to do it fast. You have to take care of not only the results and the business impact outcomes, but also you’ve got to keep I mind the constraint that you don't have much time to bring up those insights. Of course, we saw a great example of where you need to, as a dat scientist, keep the user experience in mind. Ultimately, people can get value out of ads.

Kirill Eremenko: You know, some of the stats that Justin shared with how many people are watching these ads to the end, and how they are serving them are incredible. That just shows that they're serving the relevant advertisements, and the relevant content to the

Show Notes: [http://www.superdatascience.com/269](http://www.superdatascience.com/269)
relevant audience. A good combination of all those things, business impact, constraint on time and user experience, all combined in the world of ad tech.

Kirill Eremenko: As usual, you can get the show notes for this episode at www.superdatascience.com/269, that's www.superdatascience.com/269, where you can get the transcript for this episode, plus any materials we mentioned on the show. And, of course, the urls to Justin's LinkedIn, and other places where you can get in touch with him. If you want to work with ViralGains, you can find them there as well. We'll include links to their website. It sounds like a very exciting company with some very bright future ahead.

Kirill Eremenko: On that note, thank you so much for being here today. Very excited about today's episode, and I can't wait to see you back here next time. Until then, happy analyzing.