EPISODE 106

[INTRODUCTION]

[0:00:10.4] IK: Hello and welcome to another episode of TWIML talk. The podcast where I interview interesting people doing interesting things and machine learning and artificial intelligence. I'm your host Sam Charrington.

Contest alert! This week, we have a jam packed intro including a new contest we're launching. Please bear with me, you don't want to miss this one. First, a bit about this week's shows. As you may know, I spend a few days at CES earlier this month.

While there, I spoke with a bunch of folks applying AI in the consumer electronics industry and I'm including you in those conversations via this series of shows. Stay tuned as we explore some of the very cool ways that machine learning and AI are being used to enhance our everyday lives.

This includes work being done at Anki who built Cosmo, the cutest little computer vision powered robot. Lighthouse, who's smart home security camera combines 3D sensing with deep, learning and NLP. Intel, who is using the single shot, multi-box image detection algorithm to personalize video feeds for the Ferrari challenge, North America.

Firstbeat, a company whose machine learning algorithms analyze your heartbeat data to provide personalized insights into stress, exercise and sleep patterns. Reality AI and Quito who have partnered to bring machine learning based adaptive driving beams or automatically adjusting high beams to the US. Last but not least, aerial.ai who apply sophisticated analytics to WiFi signals to enable some really interesting home automation and healthcare applications.

Now, as of six amazing interviews wasn't enough, a few of these companies have been so kind ass to provide us with products for you, the TWIML community. In keeping with the theme of this series, our contest will be a little different this time. To enter, we want to hear from you about the role AI is playing in your home and personal life and where you see it going.

Just head on over to twimlai.com/myaicontest, fire up your webcam or smart phone camera and tell us your story in two minutes or less. We'll post the videos to YouTube and the video with the most likes wins their choice of great prizes including an Anki Cosmo, a Lighthouse smart home camera and more.

Submissions will be taken until February 11th and voting will remain open until February 18th. Good luck. Before we dive in to today's show, I'd like to thank our friends at Intel AI for their continued support of this podcast. Intel was extremely active at this year's CES with a bunch of AI autonomous driving and VR related announcements. One of the more interesting partnerships they announced was a collaboration with the Ferrari challenge north America race series.

Along with the folks at Ferrari challenge, intel AI aspires to make the race viewing experience more personalized by using deep computer vision to detect and monitor individual race cars via camera feeds and allow viewers to choose the specific car's feeds that they'd like to watch.

Look for my conversation with Intel's Andy Keller and Amil Chindiki earlier in this series for an in depth discussion about this project and be sure to visit ai.intel.com where you'll find Andy's technical blog post on the topic. Now, about today's show.

In this episode, I'm joined by Ilkka Korhonen, Vice President of Technology at Firstbeat. A company who's algorithms are embedded in fitness watches from companies like Garmin and Suunto and which use your heartbeat data to offer personalized insights into stress, fitness, recovery and sleep patterns.

We cover a ton about Firstbeat in the conversation, including how they transform the sensor readings into more actionable data, their use of a digital physiological model of the human body, how they use sensor data to identify and predict physiological changes within the body and some of the opportunities that Firstbeat has to further apply machine learning in the future.

Now, on to the show.

[INTERVIEW]

[0:04:30.7] SC: All right everyone, I am on the line with Illkka Korhonen. Illkka is Vice President of Technology at Firstbeat. A company that I first met with at CES, in fact, I met Illka at CES and he was kind enough to agree to jump on the line for an interview. Illka, welcome to This Week in Machine Learning and AI.

[0:04:53.0] IK: Thank you very much.

[0:04:55.3] SC: Why don't we get started by having you tell us a little bit about your background?

[0:04:59.2] IK: Yeah, my background is that I've been doing biomechanical engineering for basically entire professional carrier. I did, on the 90's, I was working on biomedical engineering and heart rate variable monitoring believe it or not and activity tracking but the technologies were quite different at that time.

Gradually, I did research at that time but then gradually moving towards wellness and preventative care, while finally in a later – in 2009, I decided to take a career move and went to industry first for Nokia and later, span off – with some colleagues, a company called Pulse On and with optical, one of the primary optical heart rate tracking. And later on about a year ago, I quit that now I'm with Firstbeat on doing the next level of the stacks or to say the heart rate analytics rather than measurements.

That's in brief, my background.

[0:06:16.6] SC: For those who aren't familiar with Firstbeat, maybe a little bit more about what that company does, I happen to know a tiny bit about it because Firstbeat's embedded in the fitness watch today used. The Garmin Vivo Active 3.

[0:06:33.0] IK: Yeah, that's correct. Our core business is that we are doing heart rate analytics, we transform the heart rate reading into something actionable and meaningful for the user. We were established in 2002 as a spinoff of the University of Uvascula and we first looked in heart rate reliability as a prompt or a method for controlling the athletes over training.

Soon, we realized that actually while we try to detect and prevent training with heart rate reliability based methods, those are also applicable for regular people who are struggling with stress and recovery.

That's where we started and gradually, we have been moving into other domains from elite sports. We still do have this elite sports, so we have – we are serving elite teams like NHL, NBA, et cetera kind of teams for, and their coaches, for optimizing the training and recovering of the athletes but more I think better known area where we are is that we do develop and license the heart rate analytics algorithms for wearables.

For example, Garmin, Huawei, Suunto and some other brands they are using our analytics to transform the sensor reading into more actionable information like fitness level and power consumption, stress and recovery metrics etcetera.

[0:08:17.2] SC: Now, I've had a couple of conversations on the podcast about IOT and some of the challenges of analytics and in IOT environment and I guess when I think about this device on my wrist, it's kind of like an IOT in the small and all of the challenges, well not all of them but some of the challenges that are typical for IOT, in particular the noisiness of the sensor readings and things like that. I get the impression that there's a lot of that going on with regards to kind of the heart rate sensor in the wearables.

[0:08:54.1] IK: That's very correct, that's absolutely true. The Jupiter, limitations of requirement of the IOT are that the sensor is very power limited and capacity or resource limited environment in that sense that if you don't have the super computer power there, you have limited amount of battery you want to use as efficiently as possible to have the lifetime as long as possible.

Still you want to have rich information from multiple sensors so for example, typical Garmin sports what has the motion sensor and barometer and decent to heart rate monitor and GPS and then of course all these radio. It's quite interesting from an 80's or 90's perspective, in fact, it would be a super computer.

Things with what are being done there are quite fascinating. The other thing I think which is very different from for example, as I told, I started with biomedical engineering on 90's when we were working with hospital systems and monitoring of a patient in a hospital.

There were – the monitoring was continues and controlled in a way but here, it's very liberal or there are a lot of degrees of freedom in the sensory use also that it's people are wearing the sensor and not continuously, on, sometimes they're wearing, sometimes they're not and also the way how they wear it, how they attach the sensor to their wrist and how snug it fits to the wrist and so on. They are not so well controlled and that causes a lot of variation in the signal quality also ambient progress like temperature, ambient light and motion and so on. They affect a lot of the sensor reading. It's quite challenging data what is measured by the sensor. They are quite sophisticated algorithms and detection technologies which are required to make sense out of that.

[0:11:21.5] SC: For the devices that Firstbeat is embedded in, are you pulling data directly off of the sensors and kind of responsible for giving the watch, kind of, everything from the ground up, they heart rate itself plus the higher level metrics that you're providing? Or does the watch or the sensor provide some of those metrics and you're doing higher level analytics on top of that?

[0:11:52.9] IK: It's rather the latter. Our interface to the sensor is typically the heart rate and beat to beat heart rate and also we use motion and GPS speed and things like that but we don't' do the heart rate detection but we do, as I said, we transform the heart rate and heart rate reliability data into actionable information.

It's a little bit higher level analytics based on the physiological model of how the human physiology works and how the autonomic nervous system controls the heart rate and the heart rate variability and how it can be – how this kind of data can be interpreted in terms of physiology and behavior.

[0:12:46.0] SC: Okay, before we get in to the company's approach to the analytics, maybe take a moment to talk about heart rate variability and how that's calculated and why that number is important?

[0:13:03.4] IK: Yeah. Heart rate variability refers to beat to beat variation in turbo between the heartbeats. Each successive heartbeat is having serving time instant in between them and this time difference between successive heartbeats is varying as a function of time.

This variation is related to the functioning of the so called autonomic nervous system which controls all our autonomic body functions like respiration, perspiration and plot brace and heart rate, among others. This autonomic nervous system is usually divided into two branches which are sympathetic comparison cybernetic branch. This sympathetic branch is responsible for so called fight or flight response. Whenever our physiological resources are needed, the sooner that the nervous system activates and prepares the body for the new requirements like increased energy expenditure or increased physical activity or increased mental activity.

That is actually what happens when we have a body – physical activity or when we have mental activity or elevated stress for example, it's the sympathetic nervous systems which activates. That can be seen as an elevated heart rate and in broad terms, reduced amount of heart rate variability.

The parasympathetic, the other nervous branch and that is responsible for cooling down the body after the requirements or immediate requirements of the situation are over. So to basically recover the resources and reduce the heart rate, increase the heart rate for variability and so on.

This is in broad terms, the functioning of these autonomic nervous system branches are seen in the heart rate and heart rate variability. It sounds simple and the principle is quite simple but the challenge is that of course, there's huge inter-individual and also intra-individual variance between in these – especially heart rate reliability, they are affected by several factors like different substances and in addition stress and physical or mental requirements also our metabolism and so on play a role there and people are different. Their interpretation is quite challenging.

In fact, this phenomenon was identified years ago and there's been active research in the heart rate reliability since early 70s but only quite recently during last 10 years or so, we have learned how to truly manage all these different factors which effective phenomenon and how we can transform the information into something which is useful.

[0:16:32.5] SC: In order to do that, it requires an approach analytics that takes into account not just kind of the raw data that you're pulling off of the sensors but also a model of how the body works, is that right?

[0:16:50.2] IK: Yeah, that's right. For example, our upgrade for the heart rate reliability analytics is that we have constructed a digital physiological model of the human body or body functions so we have figured out the physiological mechanisms for example, how respiration, physical activity, heart rate, mental activity, how they affect each other, how they are related and how heart rate variabilities reflects these activities.

Based on this physiological model, we can integrate the data. That's the physiology part but then we used the data and data driven machine learning basically to the data and to be or to be relevant for its individual. And we have in fact a database of 250,000 individuals where we have a wide range of people with different age, both genders, different physical fitness and different BMI or weight and height.

By using that data measured during daily life in different conditions, we can numerically adapt this model so that it fits for its individual and produces, kind of, a personalized but yet standardized outputs.

[0:18:34.6] SC: Can you talk a little bit about the physiological model to get us started? How is that expressed? I'm imagining it's a mathematical model of some sort, what's it's kind of size, shape, level of complexity and is this something that ultimately lives in the device that the end user's wearing?

[0:18:57.4] IK: Yeah, if I start from the end of the question. So yes, it realized and then we have implemented this model into a software library and indeed it is embedded into these wearable hack devices. So today, super computers in a way and how it works is basically that for each time instant we have a first level instant. We have an active segmentation of the data that we segment the data and after that, we classify and calculate certain features of the data and these physiological motivated features based on the physiological model.

Like for example, we transform the heart rate and heart rate drive within data into oxygen consumption. A moment rail oxygen consumption, we O2, or be O2 and then we detect based on the heart rate variability. We detect respiratory rate and based on these – basically these two parameters describe the level of physical activity and also respiratory rate is related to also the physical activity and based on that we can classify whether a certain time instant is physical activity or a mental activity.

And after that we can also – while looking at the heart drive reliability and whether there is sympathetic or parasympathetic dominance, we can classify each time instant whether about this more a physic or mental stress or physiological recovery and accumulating from that we can calculate different kinds of things like the level of stress and during the day and we are among the recovery. From the VO2 levels, we also can calculate energy expenditure, etcetera.

And when we combine issues like speed or a physical effort which is measured by the motion sensor, we can also calculate and using the information like personal heart rate maximum and age and gender and so on. We can calculate and estimate the VO2 max which is the fitness level of the individual and things like that. So there is actually – we have certain white papers describing that but there is actually a multitude of different physiological mechanisms embedded.

I think it doesn't make sense to go into details but just to make as an example based that we have modeled that when you start a physical activity, there is certain – the heart rate increase is not immediate but gradual and it also depends on your previous physiological state like for example, an interval training, we know that previous internal effects to the next one and we can have this kind of on off kind of off mechanisms in the model and those are taken into account. So it is kind of state compartment model heavily based on the physiology which is behind there.

[0:22:33.3] SC: And so you mentioned that there is a physiological model and then a more of a data driven model. All of the metrics that you just mentioned, are those all coming out of the physiological model before you even get to the data driven piece?

[0:22:48.8] IK: How it works is basically that we have the principles are physiological but the parameters or how it really works, how the in good data is processed or how the classification and quantification of the different physiological parameters caused, that is heavily data driven. So we have used the peak amount of data, what we have to 200 parameters of the model and tuned it also or to adapt it to each individual. So that let's say that a data driven personalization of the model which is down there. I don't know if that makes sense but –

[0:23:39.3] SC: Well in what way is it how personalized like is this machine learning type of a model that is running on the device itself or is it you are doing analytics using some type of machine learning that spits out some number of different models or you can choose between models on the watch? How personalized are you able to go with it?

[0:24:11.7] IK: We are using both approach. So when you take a new device into use, you are asked to provide your background parameters like gender, weight, age, height and activity level and those are used to put you broadly to the scale that was typical for an individual of your age and gender and fitness level but then – and that is kind of a machine learning or data base driven learning with what we continuously update based on the data what we accumulate from our users.

But then it's also an adaptive model in the sense that we use that we have algorithms which adapt the computation to your daily physiological or your normal physiological and physical activity levels and heart rate variability levels and so basically, the devices get more accurate after a couple of days of use when they have – the algorithms have learned what is your typical level of heart rate reliability and what is your typical responses and your real true heart rate minimums and maximums during different daily activities.

So the model adapts to the user and that is why also these devices are typically designed for personal use if you would for example share these wearable devices with some of your colleagues or your wife. That would be so optimal in the sense that the learning would not converge.

[0:26:03.0] SC: The learning that we are talking about here it's – I don't know it puts – what is the right question? We're not talking about training in the sense of training a machine learning model that's certainly not happening on the wearable devices, more like tracking the individual's data and adapting the model to the model overtime.

[0:26:28.6] IK: Yeah, that is probably more accurate description of that, yes. So there is in the physiological model which is being broadly optimized for your age and gender and weight and etcetera so that is being more fine-tuned to your exact characteristics based on the adaptation there but it is correct. It is more like adaptation.

[0:27:01.0] SC: And you mentioned classifying a few times. So you are ending up having to do things like outlier detection and things like that where you are getting rid of noisy day from the sensor itself or does that happen at a lower level?

[0:27:19.6] IK: We hope that most of that happens in a lower level but in fact, we have a lot of that happening in each layer as well. So especially with optical sensors there is a lot of noise, a lot of full data and in fact we have built quite a lot of intelligence into a separational detectional core and appropriate quality data and what we actually do in the model is that when we detect that input data is not valid, we have zero mechanisms to breach those gaps or survive without providing false outputs.

In fact to an extent that for example, our stress and recovery announces is still providing quite valid outputs up to the error level of 50 persons which is quite high. So as long as more data which is coming in is reflecting the real heart rate then we can provide reason of analytics. Of course when there is more false data than real data then there's little one could do but.

[0:28:42.8] SC: Right and so we have talked primarily about the things that are happening on the device itself but you've also got this warehouse so to speak of data from individuals. In addition to the work that you are doing to develop the models for the device, are there other things that the company is doing with machine learning and analytics to take advantage of that data?

[0:29:16.4] IK: Yeah, so we have two other business and one is we provide corporate wellness service called a lifestyle assessment and the idea is that we basically sell for corporates who are interested to improve or promote their employee wellbeing, an assessment service where employees get a heart monitor, a special heart rate monitor for three days. It is based on Easy Cheese so it is very accurate to one millisecond time resolution.

And they do a three day heart rate reliability recording together with the diary and by using those at those heart data and the diary data, we construct the report and give feedback to individuals about their physical activities, sleep, stress and recovery and fitness level and things like that and those are also compared to the recommendations like physical activity recommendations, sleep recommendations and what is normal at their age and gender.

And that is used very often in these kind of wellness campaigns where people are motivated to improve their lifestyles and improve their stress management skills. So it can be used as an initial message to assess the starting point and identify the exact activities they should do to improve their lifestyles and that is something that we sell as a service and that is actually one source where we get a lot of data. We did last year about 50,000 stat assessments.

Mostly in Europe, in Finland, UK, Germany, Sweden now in Singapore as well. Not so much in US yet but that is actually a good source for data for us because we get this kind of – last year we got 150,000 new data set with recommendation and background information and diary and that is actually what we learned use for the optimization and further data allotment of our technology which we used then for consumer devices.

So yeah, that is an interesting area but that's not to say that this service is actually used and they have a lot on its own. It is not just for data collection but it is really for corporate wellness business and we see that there are a lot of interest there, about 95% of the people who make the assessment find it very, very useful and they recommend it for the others. So we are looking at how to scale it up and how to make it more broadly available at the moment.

[0:32:33.5] SC: How does a company organize to tackle analytics types of problems or challenges, given the strong domain expertise required but also the analytical requirements for machine learning and traditional analytics?

[0:32:54.9] IK: Yeah, so here in Finland we have a team of data analysts experts and then we have a team of physiologist and in fact, we have organized it so that these two teams are working as a single bigger team. So they work on a daily basis with each other. So they are people who are a pro in data science and they are sports physiologist and physiologists who are experts in the physiology and that's the big thing.

What we are doing that that they are not – they are basically never doing a single project that they are the only one side of that expertise would be present. Then of course we have engineers with software background and so to transform all of that into something which can be delivered to customers or implemented into our services but the core thing there is that the maturity of that RND or intellectual work is done by these data scientists or physiologists.

[0:34:12.6] SC: I guess I am curious about the opportunities that are created as we advance, we continue to advance in our knowledge of applying machine learning and even some of the newer techniques like deep learning that might not necessarily be applicable to running on a watch but do you have any perspectives on what opportunities these technologies and the increase focus on these technologies creates for Firstbeat?

[0:34:49.9] IK: Yeah, certainly. I think we are just getting there that even though we have data from roughly 250,000 individuals in our database, it is still – if only today I think it starts to be in the scale where I would really truly call it as a big data. Although I do recognize that it's probably quite unique, a resource that only a few organizations would have or if anybody would have that amount of heart rate variability and background data available.

So we are actually applying it at the moment. We are looking at things like how we could apply more machine learning type of things and we have done some experimentation but I have to say that to date, our perspective is very much that there are challenges with applying just pure machine learning without having some kind of predominant knowledge to constrain these kind of machine learning methods. There is a risk that you over optimize the certain data sets.

And data generalized and I think one of the learnings, what we have done and why our methods also work in the wearable environment and why they work in almost all of the individuals is that the data are based on these kind of physiological principles and we used only data to kind of optimize those principles and make them numerically work and in that sense, we are not very heavily using for our heart rate data. We are not using this kind of pure machine learning like deep learning kind of methods.

We have and we are experimenting on that but I am actually not a great believer that that would be the optimal way to go.

[0:36:58.1] SC: Great, well Ilkka this is a really fascinating conversation and I appreciate the way that you're combining the physiological models with the data driven models and I think clearly the domain knowledge is important here as is the case in many if not all of the used cases that folks are having success with. So thanks for taking the time up.

[0:37:26.8] IK: Thank you very much. It has been good for me as well.

[END OF INTERVIEW]

[0:37:34.0] SC: All right everyone, that's our show for today. Thanks so much for listening and for your continued feedback and support. Remember, for your chance to win in our AI at home giveaway, head on over to twimlai.com/myaicontest for complete details. For more information on Ilkka, Firstbeat or any of the topics covered in this episode, head on over to twimlai.com/talk/106.

Thanks once again to Intel AI for their sponsorship of this series. To learn more about their partnership with Ferrari North America Challenge and other things they've been up to, visit ai.intel.com. Of course, we'd be delighted to hear from you either via a comment on the shownotes page or via Twitter. Directly to me at @samcharington or to the show at @twimlai.

Thanks once again for listening and catch you next time.

[END]

