Gil Tayar: Not Only Cars: “AI, Please Test My Apps

Gil Tayar: Hello, everybody. My name is Gil Tayar. I'm from Applitools. I'm a developer. I've always been a developer, 30 years of experience. I'm going all the way back to the '80s. Done everything possible, been a developer and architect, founder of a company back in the days, QA manager, document writer, whatever. Currently testing the code I write is my passion. I test everything and that is what led me to Applitools because one of the things I could never do was test how the application looks like and not just how it functions. Applitools creates tools exactly for that. For me, it's in a way the missing piece of the puzzle of testing, testing how the application looks like and that is what Applitools does. We deliver visual testing tools. If you're interested in testing how your application looks like and not just how it functions, feel free to contact us.

We're going to talk about AI and testing. This is what my talk is about and talk about AI in cars and see the parallels between them. First of all, what I want to say is very simple. Do not panic. I've been talking to a lot of testers about AI and I get the feeling that there is some slight worry about losing jobs for AI and the TLDR for this talk is don't panic. If you get one thing from this talk is do not panic. AIs are tools. They will help you in your job. They will not make you lose your job. Hopefully, I will be able to convince you by the end of this talk.

Let's first talk about driverless cars. In the 1st of May, 2012, Google passes Nevada's self-driving test where they had to, I think, drive on a highway or something without the driver touching the wheel. Obviously there was a driver there just in case, but the car was driving itself. I remember being amazed at the fact that there is such a thing in 2012. It was for me like science fiction. For me it was this should not be happening. AI should not be that good, not 2012, not now. Maybe in 10 years but definitely not now, but it did happen.

Google proved to the world that AI is serious, serious business, not only to me, more importantly, a lot of people were shocked by this. All the car manufacturers suddenly basically went to war. Ford, GM, Mercedes, they understood that their future is in driverless cars. Driverless cars, autonomous cars are mostly software. AI is software. Google, Apple, Intel, Uber, Tesla and others are also in that same war to see who dominates that space.

Why is everybody so interested in driverless cars? The reason is very, very, very simple. Driving is boring. It's boring and it's stressful. That hourly commute in the morning, hopefully just an hour, is very, very boring and very, very stressful. Everybody knows that long-term nobody wants to drive their car. If they get a driverless car, they will use it. People will use driving cars because for most of us it's not a good thing to do. Yes, there are fringes that enjoy driving. That's fine. But most of us prefers sleeping in the back.

Because driving is boring, it's also dangerous. If autonomous cars can avoid the accidents that we see happening because of the boring parts, then we're also saving billions, possibly even trillions of dollars in that. That is a huge, huge, huge amount of money.

What happens if we let driverless cars drive us? Well, we'll start concentrating on the interesting part of the journey, working during that car, doing the really interesting part, figuring out where we want to go, what we want to do when we get there and not so much ... We'll start concentrating on the what, not so much on the how. That is for me is just like testing. Testing is amazing.

Figuring out how to ensure the quality of an application is very, very difficult. There are two parts of that difficult. There's the interesting part which is figuring what to test, what the infrastructure needed to test that, what to test, what not to test, et cetera, et cetera, the thinking part. There's the boring part, another page, another form filled, another form filled to check that must ... They only accept numerics and only have this length or that length. That is the boring part.

In this talk, we'll talk about how autonomous cars take out the boring part, leave us with a more interesting part and we'll talk about how AI in testing takes out the boring part and leaves us with a more interesting part.

As we said, driverless cars are here so let's talk about autonomous driving. We're already talking about millions of miles driven by driverless cars. Google already has three million miles in its belt of driverless driving and Tesla has, with its autopilot that is working right now in cars, has already about 200 million miles of driverless car and probably even more. Is Tesla winning this fight? The answer is no. Tesla is not winning because Google's car operate at a higher level of autonomy than Teslas. To understand that, we need to understand what those levels of autonomy are. There are five levels. In 2014, there was a classification that was adopted by the National Highway Traffic Safety Administration. That's a mouthful. Five levels that define autonomous cars and everybody is about those set of levels. Let's talk about them.

First of all, level zero. First level of driving is level zero. No automation whatsoever. Congratulations. You are the proud owner of an autonomous car or, in this case, a tricycle, level zero. Level one is where a lot of cars are now. We're talking about cruise control. We're talking about automatic breaking systems. This is driving assistance. The human is still under control. The human is still driving. The AI is only assisting while the human is driving.

My brother-in-law did a driving test on an automatic breaking system car and the guy is sitting next to him said ... There was a car in front and that car stopped. The driver said, "Just push on the gas. Don't worry." My brother-in-law was going like, "Oh, no. Wait, wait, wait. There's a car in front." He said, "Don't worry. The automatic breaking system will take care of it." My brother-in-law pressed on the gas and the car stopped. This is driving assistance.

What made driving assistance possible? Why didn't we get it earlier? There are two answers actually. The first answer is vision, better sensors. Once the car can see the world around it, it can act on that. The important technology is LIDAR. LIDAR is light, detection and ranging. It functions like radar. It pulses laser light that the car spreads all around it. The light comes back to the car and, based on the time passed, the car understand what is going on around it and the range of what is going on around it.

Stereoscopic cameras are also important. Some cars use that because just like our eyes that can figure out the depth because we have two eyes, so stereoscopic cameras can figure out the depth because there are two of them. Yes, some cars use radar, real radar. That's one set of technology.

The other set is better algorithms. We have just better algorithms and better computer power to use those algorithms to understand what is going on and to figure out, oh, wait, there's a car here and it's like 10 feet away. Let's push the break that hard to make the car stop.

Level one is nice but the human is still in control. The human is still steering. What about level two? Level two is partial automation. This is where Tesla's autopilot is. This is where GM's Super Cruise is. The human is still in charge but the human can, in some situations, take their hands off the driving wheel. We're talking mostly highways here. Eyes are still there monitoring and the AIs will a lot of time say, "Wait, wait, wait. This I can't handle. Please come back to the steering wheel."

What are the technological advances? Again vision is incredibly important, but the algorithms operating on that vision have advanced and it's all about machine learning. It's actually not algorithms. We'll see in a second that's not really algorithm. It's machine learning and deep learning and neural networks that drive this technological advance.

Well, first of all, let's talk about neural networks. What are they? They're actually very simple. Think of these nodes which are the neurons just like in our brain and they accept input. There are neurons that accept input. That input passes and changes through the neural networks and finally we have output. What does that help us? Let's take an example and the example is an image. The image is the input. This is a convolution on neural network where the image is an input and the output is a number like from one to four whether it's a dog, a cat, a bird or a boat. A neural network takes that image which ... What is an image? It's just pixels and pixels are numbers in that sense because they're the R and the G and the B of the colors. We get numbers as input and the output is the classification of that.

How does do neurons network? First of all, they're layers. You can see layers. In this case, there are one, two, three, four, five, six, seven layers. But how does the neurons in those layers work? The answer is very simple. They accept input and the input are numbers. Each input is multiplied by a weight and then summed up and then pass through a certain function, doesn't really matter, and that's the output. The summing based on the weights of the inputs, very, very simple.

How can such a simple thing create complex algorithms, complex understanding? Nobody really, really understand that but it does. Who decides the weight of each individual neuron? Who decides what W1, W2 and W3 in each of those neurons in the neural network happens? The answer is training. Let's understand how this works. Let's take the example of classification. We feed an image of the dog. We know it's a dog. We feed that image to the neural network. It passes through and we give random weights to each neural network, passes through and exits as a classification: dog, bird, cat, boat.

If its correct, fine. We go into the next image and we train it. But if it's not correct, we backpropagate in some mathematical way which I don't really understand. We backpropagate and we fix the weights. Then we do it again. We give another image and another image. If it's okay, we're fine. If not, we backpropagate and fix the weights. That is it. We feed thousands and hundreds of thousands of images until in essence most of our answers are correct, the outputs. This is the learning part of the machine learning. Once we have those weights, we give it an unknown image and out pops the classification. It's like magic. It's that simple.

This is why the mileage numbers are important. Google and Tesla and GM and Ford aren't just testing their driving algorithms, their driving neural networks. They are training it.

What about deep learning versus machine learning? Well, machine learning is the major category. It encompasses neural networks but also other kinds of algorithms. Deep learning is when we use our really, really big computing power that we have today with all the CPUs, the fast CPUs and the fast GPUs and we create neural networks that are deep. They have lots of neurons and they have multiple, multiple layers and thus creating better classification and better understanding of what's going on around.

How does this work in cars? The answer is there's no one answer. There are lots of applications of deep learning and machine learning in cars. One example is figuring out what the signals are, the traffic signals and the stop signals. For example, a neural network that ... In comes the image and out comes ... We have these traffic signals. There's a stop sign. There's a stop and there's a speed limit sign, et cetera, et cetera, et cetera. That's one application. The other application is the more obvious one. In comes an image, out comes the angle of the steering wheel that needs to be done to maintain the lane. Those are the two kinds. Obviously the second one is what we think of when we think of driverless cars but there are lots and lots of applications.

That is how we reach level two using machine learning and deep learning to figure out and understand what is being seen by the car. What about level three, conditional autonomy? This is much more autonomy, still the same algorithms, just much more autonomy. Humans need not monitor. But, in some situations the AI says, "I can't handle it." Let the driver take over.

Level four is where Google is. Level four, the AI totally takes over all the driving. You don't even need a steering wheel. In this case, we see a picture of a Google car. No steering wheel. The caveat and it's not final is that there are times and places that the

car cannot deal with: snowstorms, fog, heavy rain, et cetera, et cetera. The car can't deal with it. It doesn't see very well. Its level of autonomy is not enough.

Finally, level five, the holy grail, nobody has reached it, is where the car can drive anywhere a human can: fog, rain. It doesn't really matter. Holy grail, driverless under any condition.

Great. What about autonomous testing? We're here for autonomous testing. In my company Applitools, just hired a new guy in addition to our algorithm and AI team and we started talking about the similarities between driving and testing and brainstorming this idea of levels of autonomous testing. These are our thoughts on this. I think they're pretty well baked.

Level zero is no autonomy. It's not no automated testing. It's you have full automated regression testing but you write everything. The problems are changing the checks. Every time the application changes, you need to change the checks, the assertions. This field changed. That field changed. [inaudible 00:16:48] on all the application needs to be done on every test you have. When the test fails, it's a bit difficult to understand why it failed but you have level zero autonomy on testing.

Level one, driving assistance. Let's remember what was the technology that made level one driving assistance available? The answer was all about vision. It was all about vision. Let's see how visuals can help us in autonomous testing. How does a test work usually? Test works like this. All automated testing works like this. I do an action like click. I do an action like entering text. I do another action, et cetera. Then I check. I click on the button and check that the result is what I should expect and then another action, another action, validation, another action, another action, validation. But what if the AI can see the page? Can it assist in validation? What if I give our AI vision? Will that help? The answer is yes because when we do a check, for example, for a form filled, we check the username and the email address and the password and confirming the password and whatever. All the field we need to check then and it's boring. If the AI can just see the page and understand what is going on there, then it can help us.

Level one is not so much about understanding but visual. What we do in level one AI is we take a screenshot of the page and the AI compares it against the correct baseline image and finds out what is different between the baseline which we know is correct and the new image and can find bugs very easily.

Comparing pages visually doesn't seem to be AI. Unfortunately, comparing pages is not just pixel comparison. It's complex algorithms and figuring out even if some pixel changes are correct, saying, "Yeah, it's still the same image." It doesn't matter if it's different browsers or different operating systems and different GPUs. It's still the same image. Complex algorithms are involved. Unfortunately they sometimes get things wrong because they're just algorithms. We need a better understanding of what we see.

This bring us to level two. In level two, what was the differentiator? The answer is machine learning. Machine learning is understanding better the vision of the those

pages. Machine learning enhancements help us exactly do that, help us better compare the pages one to another, not just better algorithms, but understanding the way machine learning is. Taking that image or actually, in this case, two images and figuring out what is the difference. Not only that, if she sees two different pages but the same change is in them, machine learning can figure out that it's the same change. If we have a sweeping change across lots of pages, AI will help us understand that it's the same page, group them into one and just let us validate or invalidate that as a bot.

Level three, conditional automation. What do we get there? We're getting better machine learning. Remember in level two, we validate it against baseline. Who created that baseline? Well, we did. We looked at the page and the first test we said, "Yes, this is a good page." But what if the machine could understand that a page is good or not good without a baseline? What if we fed it with lots and lots of pages that are good and tell them, "These are good pages," and use machine learning to explain that these are good images and feed lots and lots of pages with bugs like in the sample page that we're seeing here, bugs in how the page looks, bugs in the values of the fields.

What if we could feed those good pages and bad pages, tell the machine learning algorithm that these are good or bad pages and let it figure out for the remaining pages? In this case, we wouldn't need a baseline or at least we would only need superficial validation of a baseline by a human. This is level three, checking any page for design and data problems without a baseline. As I said, we use machine learning to do that.

But human still drives the tests. Human still say, "Do this action. Do this action." We need a human to understand the application itself to drive the test to figure out what to test and the AI assists us in determining whether the page that derives from those actions is good. But the person writing the test and driving the test is still the human.

Level four, partial automation. We're seeing here an example where an AI is running that green Pong character. It's understanding how to play a game of Pong. Nobody told it how to play a game of Pong and yet it does through machine learning. It's driving the game and not only understanding the game. Level four, partial automation, the AI observes users using the application, understands what are good results? What are bad results? What are the actions that lead to good results? What are the actions that users do? Once it has all that data and lots and lots of data, it can start driving the test itself, observes interactions over time and start driving and generating the test. This is called reinforcement learning, by the way.

Level five is science fiction. Level five is the AI talking to the product manager, understanding what the product manager needs from the application and then driving the test itself. Why am I saying that this is science fiction? Well, very, very simple. Machine learning is very good at giving an output of numbers or images, get an output of whatever classification or whatever we need. Understanding is way beyond that. AI cannot really understand logic, cannot really understand these things. We're really in level of AI here and frankly we can't do the job of understanding the product manager and creating the test from it so how can we expect an AI to do that? AI that does that

needs to be in essence much, much, much smarter than humans to really, really do that. Yup, that's flying on the wings of fantasies.

I added a level, six. This is Skynet. This is where the robots and the AI takeover. There are no more human beings so there are no more applications to test which brings to me to the philosophical questions of who tests the robot themselves? We're way off tracking.

Are we there yet? Should we go looking for another job and start worrying? Answer these questions for me please. Does your job require thinking or is it rote and repetitive? Are you just creating more and more tests or are you thinking about those tests? If you're thinking, you're good. AI will not replace thinking. It won't be replacing you. No, that's wrong. It will not be replacing you. It will be helping you. Otherwise, if your job does not require thinking, if it a tool took away the repetitive and boring parts, would you be doing thinking work? The answer is, if you will, if a tool like AI just did the boring parts, will you be thinking more? If you will be thinking more, you're good. AI will be that tool. It will be helping you.

Otherwise, can you start thinking and not just doing repetitive and boring things? Can you start thinking about what you're doing? Probably the answer is yes. If you can, you're good. AI will not be replacing you. It will be helping you. Otherwise, yup, if you can't think, if all you're doing is boring and repetitive. In 5, 10, 15, 20 years, yes. AI will have replaced all the boring, repetitive jobs you will apply to.

Remember, AI is a tool. Tractors haven't replaced farmers. Farmers need to do all the important stuff. Tractors have helped farmers be more productive. AIs will not replace you. AIs will help you be more productive. Use AIs as they come. Benefit from the AIs and don't panic. Thank you very, very much and have a good day.