Chris Peerman (00:00):

Good afternoon, everyone. And welcome to this next session of Adelaide Festival of Ideas. My name's Chris Peerman, I'm the director of forensic science essay and the chair for this next session, which is entitled My teacher said I would need maths one day. Now our speaker this afternoon is Dr. Duncan Taylor, a brilliant scientist. And I think a very good example of some of the great young minds in south Australia. Duncan is a principal scientist of forensics statistics at forensic science essay and an associate professor at Flinders university. He was the awarded the stem professional award last year as part of the essay science excellence awards for his outstanding contribution to forensic interpretation, particularly DNA evidence. Now don't is going to speak for about 35 minutes and there'll be time for questions at the conclusion of that. And as you can see Dunkin is the proud owner of a very impressive beard. And one of his hobbies is mastering new and obscure tie knots, a good example of which he has on today. Thank you, Duncan.

Speaker 2 (01:15):

[Inaudible]

Duncan Taylor (01:16):

Thank you very much for that introduction and thank you everyone for turning up today. The title of my talk is my teacher said I would need maths one day and I'm going to be speaking about a few mathematical techniques that we use in forensic science to help solve crimes. I'm going to try and break them down in understandable chunks. Now I'm aware that running concurrent to this session is a session on mapping. I just want to make sure everyone's aware that this is a no way associated with the napping workshop. All right. So the three different techniques that I'm going to talk about today Markov chain, Monte Carlo, I'm going to talk about Basie and belief networks. I'm going to talk about artificial neural networks. All of those things may sound completely foreign or, or terribly scary at the onset of this talk, but hopefully by the time we get to the end of the talk, you'll see how pretty cool they are and how we can use them in ways to help solve crime and in various other ways as well in society.

Duncan Taylor (02:20):

But before we even start talking about any of those, you might wonder to yourself, how did statistics all begin? Where did, where did it all start from her headed? We get to where we are right now. And if you were to ask yourself that question, you might go right back to the polyps Peloponnesian war in 400 BC. And so the story goes, there was a spot in general that was trying to sack a particular city and he and his army were crouched behind a rock. And they could tell that this wall was, it was a certain height or they needed to estimate how high this wall was an exposed on. Part of the wall was, was the stonework. And if you had enough time, you will be able to count the number of bricks in that exposed part, and then estimate from that how high the wall was.

Duncan Taylor (03:10):

But the problem here was that as soon as you got out, if you try to take any time to count the bricks, you would be shot by the people on the top of the wall. So what this general did was send his men out one by one to very quickly peek at the wall, have a guess at how many bricks they thought made up that exposed area of the wall. And then at the end, he would tally up all the guesses from all his army. And he picked the most common value. And in statistical terms, that's known as a mode, the most common value in a set of values. So you could think of this as perhaps the first use of statistics for practical use.

Perhaps perhaps you're not impressed with that particular example. So we could skip forward a bit and we could skip forward to the work of our kindie, who now were, were around 880.

Duncan Taylor (03:59):

He was known as the philosopher of the Arabs and his job was to work in what was known at the time as the house of wisdom. That would take information coming in from other lands a lot, a lot from Greece at the time. And he would translate it into Arabic for his people. And because he saw so much interesting information coming into his country, he became quite prolific in a number of areas from mathematics and statistics to also weather prediction and perfume production, and all, all sorts of things. But one of the things in particular with statistics is that he came up with a CITIC statistical way of deciphering encrypted texts. Okay. So that was sort of the first use of mathematics or statistics for at decrypting a method, or we could skip even further forward into sort of 17 to 1800 Ady. And we could look at the word work of Gottfried Ekin wall who actually coined the term statistic.

Duncan Taylor (04:54):

And he was being used by the government of his country to take measurements on the population. So size of the army, the number of people, the amount of tax that people paid. And he used that to inform the government on how they should proceed. And in fact, statistics gets its name from meaning for the state. Now, of course, it wasn't long after that, that statistics were actually then used by the people to judge how well the statements use was performing rather than the other way around. So this, this is probably the history of statistics and all of these things lead up to these complex statistical techniques that we now use every day.

Duncan Taylor (05:36):

So I'm going to start off talking about Markov chain, Monte Carlo, and that's split into two different parts, the Markov chain part and the Monte-Carlo part. And I'm going to start off with a bit of ancient Russian poetry called deepen Serbia's minds, elect naught. So this poem was written by Alexander Pushkin and it sort of goes deep in, in Serbia's mind, mindset, naught subdue, your proud and patients' spirit, crushing toilet, and lofty thoughts shall not be wasted. And it goes on and on and on. And you could take a few things away from this particular poem when might be that Russians could never be accused of being an overly cheerful group of people. But you could also take away or start thinking about this poem and as you're reading through it, what is the chance that the next letter I read as a consonant or the next letter I read is of how, and if you reading this poem and you did have that thought you would being a very small select group of people that included this man that's on the screen, Andre Markov, he was interested in, in this particular problem.

Duncan Taylor (06:45):

So how can I gauge what the next letter is going to be a consonant or Val? And there's a very simple way you could do it. You could tell the apple, all the consonants or the veils in that text, and you could create a pie chart, like shown up the top of the screen there. And you could make a guess based on that pie chart based purely on the frequency of consonants and vowels in the text, that that would work to some extent, but perhaps a much better way to do it would be to look at what's known as transition probabilities. So if you're currently sitting on a continent, what's the chance that the next letter is going to be a constant. If you're currently sitting on a constant, what's the chance that the next letter is going to be a Val, and if you're don't veil, what is chance it's going to be another veil or another constant.

Duncan Taylor (07:27):

And what you can do is tele up as you chain through the letters of the poem, each of these transition probabilities, and this allows you to estimate with much better frequency with much better accuracy. The chance that you're going to land are constantly, and these Markov chains are used, then they've been used in the past. And they even used today in a whole bunch of applications. So in the past a common use of these Markov chains was weather prediction. You can imagine if you're looking up and it's sunny and it's bright and it's warm, then there's a high chance that the next day is going to be sunny and bright and warm than it would be freezing cold and rainy. So you have this simple mark of chain sort of behavior that also used webpage rankings. So yeah, each web page that has a link to another web page there's a chance that a user will click on those links and go to other web pages.

Duncan Taylor (08:19):

And just like the old adage, all roads lead to Rome, quite often, websites or links lead to one big sort of website. That's the authority on a particular topic and the transition probabilities to get to different websites, make up the Google rankings that you see when you do a search. I guess another interesting application of these things is if you ever play one of these like massive multiplayer online games, like world of Warcraft, and you're creating a character and you click randomly generate name and you end up with like snail log, the Bavarian that randomly generated name is based on a back off chain. So you start off with a random, and you have a transition probability to go to some other letter in some other letter and so on to create your name. So that's the mark of chain part. We'll. Now you got the Monte-Carlo part.

Duncan Taylor (09:05):

And this particular part of the journey takes us to 1940s in Los Alamos and a secret laboratory to do with radiation testing. And we have the man on the right of the standards, Stanislaw alum who was working on a particular problem, which had to do with radiation shooting, how much shooting was required to protect people from certain radiation. And this was quite a complicated problem, and it had little variables in it, and it was proven task two to compute at one point when he became unwell, just, just with a cold, no nothing. And he, he was, he was unwilling and in bed, he would play solitaire. And he started thinking to himself, I wonder what the probability is that, oh, we'll get a winning hand of solitaire in any one particular game. And so we started thinking through all the different permutations and combinations of cards that you could get, and the probability that those would lead to a winning hand, depending on how you played at your solitaire game.

Duncan Taylor (10:04):

And this too turned out to be very, very difficult and taxing because of the great number of combinations and permutations of cards that could be got. And he can, I came up with the brilliant idea. If I want to figure out the probability of obtaining a winning hand, what I could do is just play a hundred hands and see how often I win. Okay. This is like taking a random sampling. So if each hand is a random shuffling of the cards and you played a hundred games and you won five times, then you would think that maybe the probability of obtaining a winning hand is a 5% or 0.5 without having to do all the complicated calculations that actually figure out the exact probability of winning. And he was able to apply that to his problem for shooting during a number of simulation experiments. And he did this with the help of a man by the name of John Von Nyman, a Java nine minutes reputed, to be a genius of a man with immense cognitive abilities.

Duncan Taylor (11:03):

He's made contributions to this work with centers like Ulaan, but also computer science and economics. And he was the instigators of game theory and a number of other topics. I'll digress for a second here. There's, there's a story about someone that approaches Jonathan Nyman at a party and poses to him a mathematical problem. This mathematical problem goes along the line of if you had two bikes that were 20 kilometers part riding towards each other at 10 kilometers an hour, and then you had a fly that started on the wheel of one bike and fluid, 17 kilometers an hour to the wheel of the second bike. And then back the first and back to the second first and second, the song until it was squished in the middle help, far with the fly have flown in total.

Duncan Taylor (11:54):

So there's a difficult way you can do this, which is an infinite sum of a geometric series. And there's a very easy way you can do this. And the easy way to do it is to say, if you have two bikes, 20 kilometers apart going at 10 kilometers an hour, they hit in one hour. So for flies flying at 17 kilometers an hour, it's flowing to 17 kilometers. Okay. So John was, Jonathan Nyman was posed with this problem. He thought for a second, he went 17 kilometers. This is of course was in miles because it was a while ago, but I've updated it 17 kilometers. And I go, Hey, the guy that posted the problem said, oh, you know, the trick, Andy Johnson, what trick? I just did the infinite geometric salmon, all right.

Duncan Taylor (12:36):

Fund that you can have a maths party if you ever go. All right. So they forgot Morozov oh, that's right. So Monte Carlo, they came up with this method of, of random, random sampling to obtain probabilities. And because this was a secret facility in Los Alamos, so they couldn't, they had to give it like a code name. And so they thought of the Monte Carlo casino, which we've got pictured in the middle of there. And so this became known as the, the Monte Carlo aspect. All right. And then in the 1950s, these two things came together as Markov chain, Monte Carlo, which is what it's become very useful with solving a lot of complex problems now. And if you think about Markov chain, Monte Carlo, it's probably best thought of like a game of hot and cold, but I'm going to start off with what I'm calling here the most boring game, hot and cold in the world.

Duncan Taylor (13:21):

So you've got all these possible spaces that you could guess, and you start up at the top left there and the hotspot where you want to be used in that shaded blue area. So you can imagine if you're playing this with your friend, what you could do is you could take one tiny step forward into this next square and your friend might get hot. And then you could take one tiny step forward into the next square and your friend, like a hot, a warmer, and then warmer and warmer. And you would get all the way down to sort of in the middle here. And then as you started getting further away, your friend with your cold and you take another step in the same direction and your friend would say, I say cold, you're getting colder. Do you know how to play this game?

Duncan Taylor (14:01):

Okay. And you'd go back and forth and back and forth and back and forth to eventually get to the hotspot driving your friend insane at the same time, I imagine, but you wouldn't be deterred in your game of hot and cold. You would continue on hot and cold, hot and cold until you have the very end. And at the very end of this game of hot and cold, you would look back and you would go, that was the spot. All right. So this apart from being the most boring game of hot and cold in the world is what we might call it. A heuristic search or an exhaustive search where you test every single combination you possibly can, and then pick the best one at the end. Okay. Now, to help Marco chain Monte Carlo works a bit better, more like a normal game of hot and cold.

Duncan Taylor (14:43):

So you start somewhere randomly on the board. You make a guess it's hot, or it's cold. You either accept it or reject it, then take another guests, another guests, and you sort of get you wiggle your way into the hot zone. And eventually you end up in, in the shaded area there. And the advantage of playing hot and cold as a normal person would play hot and cold is that you've only looked at these shaded squares on the board, as opposed to the heuristic method where you've looked at every single square on the board. So you can carry out very, very complex problems, only looking at a very small fraction of the overall answers and still get to the hotspot, still get to the good spot. And this is what we do in forensic science when we're analyzing DNA profiles. So what I've shown on the, on the board here is a DNA profile from one person.

Duncan Taylor (15:36):

And I won't go into too much detail, but we look at a number of different regions and you can see perhaps that each region we have see a cluster of two little bumps next to each other, or we see a cluster of one bump. When we look at evidence samples, more than one person has contributed DNA to them quite often. And so what you get as a whole bunch of these nice looking DNA profiles or overlaid on top of each other in this big sort of jumbled mess, what we have to try and do is tank detangle that complicated series of peaks, and think about how different combinations of these simple profiles could make them up. And you can imagine that when we're looking at this profile and this happens to be from four people, we might consider that each person could be considered could be donating any way, any one of say, 10,000 different amounts of DNA.

Duncan Taylor (16:31):

And they could be degrading to any one of a hundred different gradations of degradation each and each region that we look at could be amplifying at any one of a hundred different rates of amplification efficiency. And each person could have tens or hundreds of different genotypes. And if you actually multiply all this, get together, what you end up with is a number that's to the power of sort of 80 or 90. So that's a number with 80 or 90 zeros behind it. And that is equivalent to the number of atoms in the observable universe. So if we were do extend our hot and cold game analogy, we would, when we're solving these complicated DNA profile problems, this is like playing a game of hot and cold over the entire universe where the hotspot is one atom. So it really, really complicated problems. And this system of Markov chain Monte Carlo is one that we use every day.

Duncan Taylor (17:28):

And it's used all through Australia and New Zealand and increasingly around the world to carry out these complex profile evaluations. And that allows us to compare people to these complex profiles and perhaps make comments on whether they could be contributors of DNA. All right, next topic, Basie and belief networks. And I want to start off again with a simple example. And some of you might've heard this example before it's called the Monty hall problem. So in this particular problem, you've got three doors you're on a game show behind two of the doors are goats behind one of the doors, a car, all right, you want to choose the door that has the car behind it because you would rather a car than a goat. So let's say that you come along and you choose door. Number one. Then the host of the game show comes along and he knows what door holds the car.

Duncan Taylor (18:27):

And he exposes. And now he opens up another door that you haven't chosen in this case door number two. And he shows you what to goat. And the question he puts to you is, do you want to stick with your original choice or you like to switch to the other door that hasn't been revealed yet? And you have three possible outcomes here, either. You think you should stick with your current door because you've got the best chance of winning. You think you should switch to the third door or the door that hasn't been opened, because you've got a better chance of winning. If you do that, or you think that it doesn't matter because at this point it's 50 50. So just out of interest who thinks we should stay with our current door, who thinks that we should switch to the other door, who thinks that it's 50 50, and it doesn't matter.

Duncan Taylor (19:12):

All right. So we're pretty evenly split room. As it turns out, it's always better to switch doors. You're twice as likely to win. If you switch, then if you stay where you are, and we can explain this, if we extend the Manticore game to perhaps consider a hundred doors and you pick a door, and then when you pick that door, you realize that if there's a hundred doors, you've actually got a one in a hundred chance of being right, just randomly. And it means that other 99% is on all of the other doors. So then you could think that the host comes along and eliminate eliminates 98 of those other doors leaving. Just one, that original door that you chose still has the 1% chance of holding, having the car behind it. But now all of the other 99% is concentrated on the one door that hasn't been revealed.

Duncan Taylor (20:04):

Okay. So you now actually 99 times more likely to win. If you switch, same thing happens with the three doors. It's just that it's only, it's obviously less extreme. So it's less easy to, to think about it in that way. Now I bring this up because it's a difficult mental problem. And even though it seems like a relatively easy question to ask, and we can address these sorts of difficult problems by con by using these things called Basie and belief networks. And these basic belief networks are based on the work of the Reverend Thomas base shown over here on the right of the screen there. And these are like graphical ways of showing complex formula. So you could derive the formula would prove to you that the Monte hall problem is actually as I described it, or you can create a Basie network, which sort of is these bubbles that I show on the screen here.

Duncan Taylor (21:01):

And the idea is that if you provide information to some of the bubbles, then the probability is updated in the other bubbles, and you can use that other information to a great effect. So one example of this in a forensic context would be the work of a man by the name of Colin akin, who looked at a series of 300 homicides. And it had look at different aspects of those crimes. So he looked at the sex of the victim, whether the offender was married with the offender, knew the victim where the victim was last seen in the victim's age and the method of killing. And he found out that there's various relationships between these different factors that are involved in crime. And so by having some information from the crime scene, you could then use these Beijing networks to obtain information about other aspects that you don't have information for.

Duncan Taylor (21:54):

So for example, in his work, he used this scenario. So a female victim between zero seven years found strangled in her own home. You can see that you have information about four of the seven different nodes within that base in network. And if you supply that information about the crime in to those four different nodes, then you can get more information about aspects in the nodes that you don't have any information. So by supplying this information, we can see from the table here, we've got the initial probability and the revised probability once the information has been provided. So is the offender living with a partner? Yes, no. Initially is 24%. You provide this information about the crime. It goes up to 33%. Well, it's the victim known to the offender started off at 57% known, provide this information that goes up to 66% known. Okay. So you can use this in an investigator capacity.

Duncan Taylor (23:00):

The other way you can use it is to actually provide an evaluation of evidence in court. So I've got here an example of a south Australian case. So this is a case of Mr. Drummond, where he was accused of jumping out of a van and attempting to kidnap someone off the side of the road. In this particular example, there was no evidence found of the victim on dumbest on Mr. Drummond or from Mr. Ramadan to the victim. And he was despite this still convicted. And then later that conviction was overturned. So we can use these Basie networks to evaluate the absence of evidence in light of the question being asked by the court, like, did they have contact together? And I'm not going to go through all of the different aspects of this network, but basically what I want to show you is that you can consider all these different factors.

Duncan Taylor (23:51):

Like did they have previous innocent contact? Who is it that tried to adapt to this girl? Was it Mr. Drummond? Was it someone else whose DNA was found on each other's tops? And you can supply information into this network, which is something we did in, in this particular case. And it all propagates through the network back to the node of interest. And then you can start make making comments like these findings of no DNA are twice, two times more probable if he did not attempt to kidnap the victim than if he did. And this allows the courts to use all the information in a sort of wider case context. So this is something that we are also doing in forensic science. All right, last topic I want to talk on is artificial neural networks. And you can imagine that your brain is made up of a number of these little neurons and they get some sort of input via visual audio input or sensory input of some kind.

Duncan Taylor (24:48):

And the little neurons will take that input and they'll pass it on to other neurons. And those are the neurons will have a certain activation energy. And then they'll pass that signal into other neurons. And eventually a thought is formed or some, whether it be subconscious or conscious, some, some thought is generated. What we're interested in artificial neural networks or artificial intelligence is trying to mimic this process, but in a statistical artificial way. And it's, there was a little work that was done pioneered by Alan Turing in this sort of biological computer type context. And today that's evidenced by the fact that we have the cheering test where if you don't know what that is. Yeah. It's a competition where a bunch of judges sit down and they type on a computer and they don't know whether they're talking to on the other end, a person or a AI artificial intelligence. And if they can't tell after a certain amount of questions, whether they're talking to a person or a computer, then you've passed the cheering test. Okay. If they can tell that it's a computer, then it's failed. The AI has failed the cheering test.

Duncan Taylor (25:55):

And when artificial intelligence was first being worked on in sort of the 1950s, they started doing so by setting up a series of big sort of what if questions? So if this happens, say if you're trying to open a door, try and use the door handle, does it turn, does it not turn? If it doesn't turn? Is there a doorbell or is there not adult doorbell? If there's no doorbell knock on the door, if there is a doorbell push. So you get these questions that split off and split off and split off in trees, but these quickly failed to, because you can't create enough questions to cover real life. There's always an infinite number of possibilities. And as soon as one of these artificial intelligence systems comes across something it's very slightly unused to it. It fails. So this then born the idea of machine learning.

Duncan Taylor (26:42):

So machines teaching themselves, and the way that this has done in artificial neural networks is creating a statistical neuron, which you see up the left of the screen there, and just like a normal brain neuron. It has number of inputs, and it does something statistical in the neuron there. And it fires out an output. And you stack lots of these statistical neurons together into an artificial brain or artificial neural network, which you see on the right of the screen there. And you supply this artificial network with some sort of input, whether it be vision or data, audio, pictures, whatever you like, you tell it, something about that input that you want it to learn. So a classic is you provided with lots of pictures from the internet and you tell it which one has we have cats in it, or which one of the pictures of cats, the neurons will learn to recognize cats in pictures from any other picture that you might provide it.

Duncan Taylor (27:35):

And the great power of these things is then you can supply a picture that it hasn't seen before, but it's different to anything you've seen before. And it has learnt features from other pictures to identify cat or not a cat in that new picture. And these have been used for all sorts of applications, I guess in real life, you might have remembered a while ago. There was sort of a quite famous application of these machine learning algorithms. So the AlphaGo system developed by Google played this game of go against the grand Basta champion at the time Lisa doll and won four games out of five. This game of goal was often thought to be like the crowning achievement of artificial intelligence because of the sheer number of possibilities of games that could be played out on, on a global. And quite as an interesting aside to this Lisa was, was from Korea and they have quite a deep connection.

Duncan Taylor (28:35):

And with this game go, there's actually university courses dedicated to it sort of considered to be reflective of a person's life and their soul and how they conduct themselves. And so when this AlphaGo system beat the grand master champion, there was quite a bit of despair and upset in Korea at the time, because it was sort of this machine, this soldiers machine had beaten the Grandmaster at this soulful very human game. And that was then followed almost immediately after, by a huge spike in Korea of enrollments, in artificial intelligence courses artificial intelligence used for driverless cars. So you can imagine a person driving around in a car and teaching the car when it should turn and how it should stay in lanes and what it should look out for and pedestrians, and to recognize the back of other cars and keep a safe distance.

Duncan Taylor (29:26):

And as these neural networks learn, they, they learn how to drive the cars better. And these were SU becoming into regular, regular use community. This is really cool. The other thing you can do is supply pictures of famous artists and your networks, and tell them to learn the style of that artist. And then you can supply a non-related picture until the neural network to apply that style that it's learned from the famous artist to this unrelated picture. So on the left of this slide here, we just have a random photo of some buildings. In the middle, we have the classic starry night per Vango. The neural networks been told to learn the style of Vango from starry night and apply it to this image. And so what it spat out on the right there is its interpretation of the image in the style of starry night, but perhaps we'll get back to a bit more science-y type stuff. Your networks have been used to great effect in the medical field to identify cancerous or tumor cells on a slide by people giving it examples of tumorous or cancerous cells and examples of normal cells. It learns the features and the patterns in the images that help it identify these tumor cells. And eventually after a while, a number of studies have shown that these artificial intelligence systems can screen slides with a higher accuracy than train doctors. How much time do I have left?

Speaker 2 (30:53):

Great.

Duncan Taylor (30:55):

Now there's a lot of automation that's going on at the moment with using computers and machine learning, and it's going to affect a lot of industries. I'd like showing this slide for a couple of reasons. One it's, it's just interesting as sort of the prediction of automation, how jobs will be affected by this sort of automation. What I particularly like showing, because I work in the biology department is that biology is pretty safe. And I show that to great, with great glean to my chemistry friends in the same building, because they're less safe. And then both of us together show that with great leap glee to our law colleagues who are very unsafe, but of course this slide has to be taken in context. So this doesn't mean that jobs will be disappearing. It simply means that the jobs will be modified in some way.

Duncan Taylor (31:43):

So for example, the law jobs, the artificial intelligence and the machine learning is going to be used to great effect to scour through previous rulings and precedents and great volumes of information that currently poor sort of legal secretaries are forced to spend hours of their time doing. And they'll be able to do that much quicker than, and much more accurately than those legal secretaries might be able to. But of course, we'll still need lawyers as unfortunate as that might sound. And this next slide is also interesting because it shows, if you ask someone, will robots take jobs over the next 50 years, you have a predominant. Yes, they will. But if you ask the slightly different question, will robots take over my job in the next 50 years? The answer is predominantly no. So this leads to the effect of robots. We're still every job, but yours.

Duncan Taylor (32:37):

It's just an interesting perception of how it's heading. All right. It though how we're using artificial neural networks in forensic science and for this psych step back to the DNA profile as much like the one that you saw before, and what we're doing is applying these artificial neural networks to classify different areas of the DNA profile. So without going into too much explanation, we have a number of different types of fluorescents that we see in a, they listed up on the right-hand side of the screen, their baseline pull up a Leo. This is just different types of, of the residents. And depending on what you look at in the profile and the pattern of data around what you're looking at you will at the moment manually as as a DNA profile reader, click on a region and either accept it as important or remove it as noise or artifact.

Duncan Taylor (33:30):

And you'll try to mirror this between two different people and they'll compare and resolve differences. This can be quite a long laborious process. Whereas artificial neural networks can be read or fed in the same exact data. And because it's all about pattern recognition, it can learn to recognize patterns in the data and categorize these fluorescents types basketball quickly, much more accurately than, than people can. So if we take an example of 18 a profile with all of the artifacts and all of the interesting information left on this is what we would see. And then we run this through an artificial intelligence system and it recognizes all of these highlighted areas is artifactual in areas that we're not interested in. And then everything left over is what we actually are interested in. And it sort of does this in a few seconds based on millions of training examples that it's learned that you've fed it in, in the previous training. Now I'm just going to end on a sort of cautionary note here. And what I'm going to do is, is start with a story. So imagine a man and his son driving along in a car along the road, and they unfortunately have a car crash and tragically, the father dies and the son is rushed to hospital. He comes into the ER room, the surgeon uncovers the Sunday. I says, I can't operate on this boy. Here's my son. All right. So how, how can this be possible?

Duncan Taylor (35:01):

Now at this point you might sort of been thinking, maybe it's an adoption thing. Maybe it's an identical twin thing or some mixed up in the ambulances. But of course there's a very simple explanation for why the surgeon would say, this is my son. It's because the surgeon is the boy's mother. Okay. And quite often you don't think about that because we have this sort of societal bias that people in positions of power and importance tend to be male. And I ring bring up this particular point because bias in artificial intelligence is quite an important topic because they learn what you teach them. If you teach them with a bias, then they learn a bias. And there's been some great examples of this exact thing occurring using machine learning in the past. So a few examples, which I've got on the board here when artificial intelligence judges, a beauty contest, white people win.

Duncan Taylor (35:56):

This is because when they have artificial intelligence systems were supplied with all the winners, all the beauty contests for all sorts of different beauty contests in the past predominantly it was white women that, that went one, those con contests. And so the artificial intelligence system learned that that's what wins a beauty contest or that's what is beautiful. Another classic example that's occurred is facial recognition. And if you teach a artificial intelligence system to recognize faces by supplying it with lots of pictures of faces and you predominantly provide it with Caucasian faces, then what will happen is we'll show him down in the bottom left-hand of the screen here, where if you've got a camera and you have all the facial recognition, it recognizes everyone except for this book guy in the middle here, another great example, we've got down the bottom. He was an experiment that was run by at one point where this avatar sort of artificial intelligence learning system that was meant to be like a chat bot was put in it's Twitter and called Tay tweets.

Duncan Taylor (37:02):

And it was sort of built on the idea that it would stimulate conversation and the more conversation that it got, the better it was doing, and it would learn that that was better. But of course, people being people when people being on social media can be a bit mischievous and unkind, they started supplying it with all sorts of inappropriate material about violence and racism. And of course it learnt that when it responded in these ways, it would get more tweets back. And so very quickly, I think it was within a day, it turned into this mega Mila maniac violence was supremacist kind of tweeting thing. It was pulled down. All right. So whatever you put into these systems is, is what you think it out. And of course you have to be careful when you are using this in our forensic science context, because you can imagine applying this perhaps to, for example, if you wanted to identify criminal behavior, you could provide a whole bunch of footage from criminals, perhaps in jail.

Duncan Taylor (37:53):

But if there was a minority group that was overrepresented in jail than just being from that minority group would start to become an important factor that would make it a sign, tell you that these people more likely criminals when it was applied to non-criminal. People I say is just have to supply the information to your neural networks in unbiased ways. And the good news is there is a lot of work that's been going into how to supply information in unbiased ways. There's even, even websites set up like the algorithm algorithmic justice league that lets you report bias in algorithms and, and gives you ways of getting rid of that bias. So just to finish up here, these three different systems, which are being in forensic science to solve crime. So you've got the artificial neural networks, artificial intelligence systems classifying that fluorescents within DNA profiles. Then you've got Markov chain, Monte Carlo solving complex problems about who might've contributed DNA to those DNA profiles. And then you've got these Basie and belief networks where we use the results of this DNA profiling in this wider case context for these very complex logic problems that the courts are interested in. So thank you for your time.

Speaker 2 (39:08):

[Inaudible].