

Intelligent displays such as this may only be a few years away thanks to the fascinating research of computer scientists who specialise in eye tracking and machine learning software. Like computerised mind-readers, eyetracking technology follows the strange patterns of our gazes, and machine learning software is used to learn what it all means. But reading our gazes is not easy. It's so hard to do that in 2005, PASCAL (a European funded network of scientists who specialise on pattern analysis, statistical modelling and machine learning) sponsored a challenge: could a computer learn to tell whether we found something useful, just by watching our eyes? Could a computer look deep into our eyes and guess our thoughts? If it could, what kind of program would it need to run?

Our eyes really do give away many of our thoughts. Only a tiny part in the middle of our vision called the fovea is capable of seeing detailed images. Everything else is just a blur. To give us the illusion that we see everything around us in perfect clarity, our eyes dart about several times a second in saccades, sampling different parts of the scene around us, and our brains glue together the separate parts to make the complete view that we think we see. Even as you read this text right now, your eyes are not sliding smoothly along as a camera might. You are hopping from word to word, often focussing on the middle of a

We thought it would be good to give the machine learning community the chance to try out their methods in a new field

word, maybe focussing twice on a longer or unfamiliar word such as *saccade*, sometimes backtracking to resample previous words, and often skipping the smaller words entirely. If a line was drawn, following the path of your gaze as you read this document, it would resemble a messy child's scribble, not the smooth line from left to right that you might have imagined.

When we look at a more complex document such as a web page or poster, it's even worse. Our eyes are flitting about the screen or paper like demented grasshoppers, and even when we fixate on something for a moment, our eyes may drift slightly, tremor or even continue to dart about in tiny micro-saccades. Not only that, but our irises also change depending on our mental state. Their main function may be to dilate depending on the light, but they also fluctuate if we're thinking hard or having an emotional response such as anger, guilt or desire. If we're thinking particularly hard or remembering something, we may look away in a particular direction, let our eyes unfocus and ignore our vision altogether.



Not a leaking pen, this is the pattern of eye fixations over a period of less than 5 seconds as a person searches up and down a Web page for the right link. Larger blobs mean the eye fixated on one spot for longer. Surprisingly few words are read. The red 'x' marks the link that was chosen. *Image produced using a Tobii X50 eye tracker, operated by Sven Laqua, Research Student, Human Centred Systems Group, UCL. Eyes belong to Peter Bentley.*

The good news is that all of the tiny, barely perceptible changes in our gazes can be measured. In 2005 a Finnish group of computer scientists did exactly that with a series of tests on 11 people. They used a Tobii eye tracker built into a computer monitor that beams near-infrared light at the pupils of the eyes to create patterns of reflections. These patterns were then used to track exactly where the people were looking on the screen, at 50 times a second. The test subjects were given a task rather like a multiple-choice questionnaire: they were presented with a question and 10 possible answers (5 wrong, 4 relevant, and 1 right) and asked to find the right answer. The patterns of the gaze of each person were then measured, as they read and reread the text on the screen. The data gathered was then used for the challenge: could a computer predict which text a person finds most relevant from only the shifting movement of his or her eyes?

### When we first started we thought there was no way it could work

“Gaze patterns contain both direct and subtle cues about users’ attention and interests, but being very noisy they require sophisticated modeling and signal processing,” according to Samuel Kaski, one of the organisers of the competition. “We thought it would be good to give the machine learning community the chance to try out their methods in a new field of application.”

Two competitions were set for European scientists: in the first, the data was preprocessed into useful, time-independent categories such as length of saccade, length of fixations, pupil diameter; in the second just the raw time-series of measurements of the eyes were provided – a much harder task.

The entries to the competitions were published in a PASCAL-sponsored workshop on Machine Learning for Implicit Feedback and User Modeling. Some attempted to use software based on finite state machines to learn to predict the child’s scribble eye gaze pattern. Others tried to assign probability distributions to the data label sequences in an approach known as conditional random fields. Fascinatingly, although each competition was won by a different group of scientists, the same kind of method came out top for both: machine learning software based on neural networks (see box).

Michael Pfeiffer and his colleagues at the Graz University of Technology, Austria, won the first challenge. They used the clever observation that, in a multiple choice exam, the answer that a person perceives as being correct is likely to be read more times, and is likely to be the last line read before the person gives their final answer. So their method ignored most of the tiny movements of the eye and concentrated on the large and conscious movements.

### Neural Networks

Unexpectedly, both winners of the challenge used neural networks to enable their computers to learn this task. Real biological neurons, such as the ones in your head, send electrical pulses to each other and are linked together in super-complex networks. Computer models that approximate this behaviour give computers the ability to learn just as we do. One of the most common models is known as the multi-layer perceptron (MLP), and it was this model that won the first competition. The MLP is a simple network of very basic “neurons”, one for each input parameter to the problem, one for each output, and one or more “hidden layers” connecting the two input and output layers. Neurons send their signals forward through the network, emitting a value as a weighted function of the values on their inputs.

MLPs are popular because they have a good mathematical foundation and they are flexible models that are easy to use. They typically use a sigmoid or hyperbolic tangent function to transform the inputs into an output, and it has been shown that a linear combination of these nonlinear functions can approximate any continuous function of one or more variables. Essentially this means that even when you have no idea how your output may be related to your input, the MLP can approximate the function that produces the output from the input. This is perfect if you’re trying to get a computer to learn something tricky, such as which series of eye movements mean a piece of text is useful, and which series mean that it is not.

The winner of the second competition used a more complex type of neural network model. Instead of using the very abstract model of neurons of MLPs, Finnish researcher Tuomas Lepola used a newer approach known as generic neuron microcircuits. This method uses more biologically realistic neurons that fire pulses at each other, and connects them in recurrent networks (the outputs feed back into the inputs) unlike the feedforward networks of MLPs. It has more connections between neurons that are closer to each other in a three-dimensional space, resulting in the formation of “circuits” that are used like a “fading memory” to represent time-series data as it is input to the network. Then the overall state of the network is read by readout functions, trained to extract the desired pattern of information. The whole idea resembles biological neural networks far more than traditional approaches, and its success at solving the challenge is perhaps fitting: a neural network that resembles our brains was best at understanding the movement of our eyes, caused by the real neural networks in our heads.