AI and machine learning for decision support in healthcare

A PRELIMINARY STUDY INVESTIGATING EXISTING SERVICES AND THE
ART OF DEVELOPERS WORKING ON MACHINE INTELLIGENCE

Innehållsförteckning

Introduction to the new edition

Artificial Intelligence (AI) and machine learning are on everyone's lips, not least after the focus it was given by the Swedish physicist Max Tegmark in book form and on TV. We often associate artificial intelligence with super-smart computers that win at chess and Jeopardy or with a changed job market where many tasks now – or in the near future – can be carried out with the help of digitisation.

But does artificial intelligence have an impact on something as tangible and human as healthcare? Of course. It is a prerequisite for the future of health and medical care. AI is a success factor, not least in prevention and early detection. By using health data strategically and systematically, we can make health-care more efficient and hopefully also cheaper. Public health can be improved.

The amount of data on each individual's health status is growing rapidly. This includes the information the patient gives during visits within the healthcare system – such as all the data that a standard blood sample provides – as well as data that the individual produces,

for example using health apps. All data on the individual that is of importance to health and is collected in a structured and systematic way is known as systematic health data.

In the right hands, the available health data could be used to provide health and medical care that is customised to each and every one of us. Unfortunately, the work of the medical care system and regulatory authorities is lagging behind, which means we cannot use our health data to the full. This is something Swelife is working on through the SWEPER project, where among other things we look at the legal, regulatory and semantic obstacles on the path to systematic health data, which leads to better use of the data – and in the longer term to better public health and a strong life science sector in Sweden.

It is now a matter of urgency. Sweden could have a competitive advantage if we quickly straighten out the issues surrounding AI, machine learning and systematic health data. We have great advantages compared to many other markets, such as social security numbers and a health and medical care system available to all.

To get full benefit from AI investments in the future, we must invest in structural changes. The changes are perhaps not as stimulating to the imagination as chess-playing computers, but they are necessary in order to really be able to harness the power of artificial intelligence and machine learning.

We must also think globally from the start and create an international connection, where we work in accordance with international standards, code systems and regulations so that we can share data across borders.

It is important that Sweden does not just become a supplier of health data, but that we build up and retain the competence to process data. This will create value for us as individuals and for society as a whole. Expertise such as this could become a new export product.

If Sweden is to continue to be a world-leading life science nation the entire life science sector must have better access to systematic health data. Researchers from both industry and academia must be able to share the data. Otherwise we will fall behind in terms of the development of new diagnostics, medical technology, medicines and new treatment methods.

Marcus Österberg and Lars Lindsköld's preliminary study was carried out with the support of Region Västra Götaland. The preliminary study is an excellent example which illustrates the challenges we face. Swelife has now published a somewhat revised edition so that the book can be widely distributed.

Lund, 22 October 2018
Peter Nordström
Programme Director, Swelife

Summary

Our intention has been to investigate the suppliers' offer as well as to find out how to proceed with the basic groundwork itself.

At this point, we must state that it is not **artificial intelligence (AI)**[1] we worked with but rather **machine learning (ML)**[2]. The people who seem to take a realistic view of the hype surrounding AI seem to agree that machine learning is a sub-division within AI in academic circles, but that we will have to wait at least a few decades before we have meaningful AI. So in the report we can perhaps relate to the hopes for AI in the future, but we want to be clear that we believe that machine learning is a much more appropriate term for the state of technology development today. Possibly even that **machine intelligence**, which is used a lot in academic contexts, sets the right expectations.

Hypothesis 1: Processing and understanding medical history and patient accounts

First, we need to distinguish between the concepts of **medical history**[3] and patient accounts. In this report, medical history refers to the disease history recorded by health professionals. It is admittedly supplied by the

patient at the time of care, but it is a controlled conversation with the aim of getting a good overall picture, as those of us in the healthcare system want to structure the history in order to provide the background to what we are doing.

A patient account, on the other hand, is told more spontaneously and in other contexts.

The reason that we need to differentiate is that in the healthcare system we have a lot of medical history connected to visits. Accounts may be recorded in an app in the form of a health diary, or in any uncontrolled way in which patients log their health. Medical history can be described as information given at a certain time and the individual's account as information which covers a period of time.

Using **NLP (Natural Language Processing)**[4], we can find out what a person is talking about, and what symptoms they have (using **NER, Named Entity Recognition**[5]). In this way we can look up medical code systems, care plans and guidelines to find an appropriate activity. During the preliminary study, we have mainly matched medical histories of the thorax sector with the code system **ICPC (International Classification of Primary Care)**[6]. However, the method can of course be used with other common code systems, e.g. **Snomed CT**[7], **ICD-10**[8] and **KVÅ**[9].

When we evaluated Amazon's AWS service for NLP, we are informed in a friendly but firm way that it does not support Swedish, and we are not confident that automated translation will not lead to the loss or distortion of information.

However, it seems to be possible to make NLP more manual (and in Swedish) through frameworks such as **NLTK (Natural Language Toolkit)[10]** and you can process what you obtained with other machine learning techniques, for example **deep learning[11].**

Conclusion

A major challenge is the lack of structured information about diagnoses, medical guidelines etc. which can be understood by a computer. If this type of information was machine-readable instead of in the form of PDF files, and computers could investigate it as human beings can investigate Wikipedia, we would be able to achieve much more. This is done using a technique called **linked data[12].** The Swedish National Board of Health and Welfare with its API "Försäkringsmedicinskt beslutsstöd" (Medical insurance decision-making support) has understood this.[13]

A promising thesis was presented at Chalmers University of Technology in the spring about summarising medical texts. We have been informed by psychiatric services in Region Västra Götaland that this would be an attractive solution since their patient records tend to be very long and difficult to read properly prior to each visit.

Hypothesis 2: Voice and conversation-based user interface

Relying solely on a Swedish voice-based conversation as an interface does not currently seem possible. Often misunderstandings arise and the question is whether we gain anything if the user must manually correct what has been said. We have mainly tested Apple SiriKit and Microsoft Azure. Azure did not understand

what we said even when we chose to read a certain line from The Godfather films, not even after three attempts, despite the fact that it knew which line it was. Admittedly the test was performed in a noisy environment, but it is certainly not an unreasonable sound image for a realistic user scenario, in our opinion.

We have also evaluated the smart speakers Google Home and Amazon Alexa in order to consider what equipment a user would be able to have at home when they need to contact the healthcare system in future. These two speakers do not speak Swedish, but for those Swedes who speak one of the world languages – such as Spanish, Cantonese or Arabic – similar gadgets may become relevant faster than for the majority of Swedes.

What we have learned from testing voice-controlled gadgets is that they are great for those who, for one reason or another, have difficulties with writing, spelling or reading, but can speak just fine.

> "A common assessment is that 5 to 8 per cent of the population in the literate part of the world has problems with reading and writing of a dyslexic nature."
>
> – The Swedish Dyslexia Association[14]

And frankly, even those of us who do not have difficulties sometimes struggle with words that we are not used to seeing in writing but know what they mean. A conversation-based or voice-controlled interface can make healthcare more accessible to people with disabilities.

Conclusion

Our conclusion is that it is beneficial to offer both voice-recognition software and a keyboard when text is to be entered. When voice-recognition is used, however, the user must have a chance to correct the text before it is sent, used or saved.

Hypothesis 3: Computer Vision and deep learning

In 2018, it seems impossible that the major cloud computing suppliers' operational services can be used for meaningful image recognition in the field of medicine. In particular, we evaluated Microsoft Azure and Amazon AWS. X-ray images of hands are described as "A white vase on a table". A man lying back on a bunk getting an injection in his arm is tagged as "person, indoor, sitting, using, woman, holding, bed, table, hand, top, green, young, white, cutting, food, man, playing", where "using" possibly explains the syringe although we would have preferred the tag "syringe".

Conclusion

But we can always work backwards and do the dirty work ourselves. That involves learning a machine the geometry behind a bad mole, or whatever you want to identify. This appears to mostly be done by creating models so that you can discover things automatically (**feature detectors**[15]) and working with neural network (probably the convolutional sort[16]). This is a possible lead for further study.

Summary

Unfortunately, we can see a number of obstacles:

- The fact that Swedish is not a priority language makes it difficult to work with NLP or to procure services. Some of the world languages spoken in Sweden may receive support before it is available for Swedish.
- The fact that medical content is not recognised in the operational solutions of major suppliers.
- Few operational products are available to procure or hire which can be used for anything meaningful. Quite often, uncertainty about liability is cited, and the issue that suppliers have still not really decided on their business model (perhaps because it is somewhat unclear where the greatest value lies).

What we can use the major cloud computing suppliers Amazon, Google, and Microsoft for is hiring computing power to train our own models in machine learning. They do not (yet) seem to have operational solutions, at least not in a medical application area.

NLP is most promising

What is most promising for continuation is decision-making support for self-care/self-triage and the creation of different information services based on the individual's symptoms. **Triage**[17] is a method for sorting and prioritising patient requirements based on their medical history, symptoms and sometimes data on heart rate, respiration, body temperature etc. In an Emergency Department, it is usually the first assessment you get before you either have to wait your turn or come first in the queue in cases of emergency.

Medical history is interesting in combination with NLP technology because we can then match the medical history with medical vocabulary such as Snomed CT, ICD-10, ICPC etc. It provides an increased structure for free text and can thus better complement other data when training a neural network.

Our opinion is that the machine learning functions available for simple use in cloud computing suppliers' services are not yet ready for the meaningful use of images. Nor is Swedish supported as a language in NLP – with the exception of IBM – where three of the nine NLP functions are also supported in Swedish.

In other words, we need to take greater technical responsibility if we are to make machine learning work – or we wait and hope that someone else will solve our problems for us. However, using the same technology that is available to the rest of the world may not put us at the top in the world of e-health by 2025. Perhaps we need to make the most of our own unique skills, as regards both technology and healthcare?

Business intelligence from the Vitalis Conference 2018

It is slightly disappointing that none of the solutions presented as AI during the e-health conference Vitalis 2018 seem to have come far enough that it is possible to use them. On the one hand, they seem to build on lots of manually structured information, in the form of triage records, or **decision trees**[18], in order to have a body of knowledge. On the other hand, they have not even used the established technology to provide the machine with contextual understanding.

An example of this is that an AI nurse suggested "fungus in the genitals" as a relevant further discussion

regarding the complaint "fungal infection of the nail". A relatively simple solution regarding the relationship between different concepts could have supported the AI nurse with the information that nail and genitals are not significantly related.

Next step – getting more people involved

On a more positive note, we have established contact with a doctor at Sahlgrenska University Hospital in Gothenburg who shares our idea of continuation and we now need to find people who are closer to care provision for a larger scale evaluation.

At the same time, we know there is a preliminary study project in Region Västra Götaland's conversion area Digital Care Services, which is aimed at self-triage. Self-triage is one feasible outcome of an investment like this. Another is an automated second opinion and quality work in general, as it is relatively easy to find discrepancies where the diagnosis is unexpected in relation to the meaning of the medical history or the lab results.

In view of what we have learned about the possibilities of deep (reinforcement) learning and the quantities of patient history we have, it seems that a combination of NLP and deep learning should be what we focus on in a larger scale project.

We hope to be able to bring together several parties' interests in a common next step, close to care provision, turning the idea of a workshop and conversion into reality.

A background to artificial intelligence (AI)

Without discussing AI as a concept at length, we use the term to refer to what is known as **weak** or **narrow AI**[19]. That is to say specialised, non-general AI. It is a machine that does one thing well but does not understand other things it has not encountered at all. You can also choose to see AI as a research area in which machine learning is the main issue. In reality, this preliminary study focuses more on machine learning.

Since AI is used as an umbrella term for a lot of things today, it is difficult to find the grain of gold in what is offered. That is why we became interested in a preliminary study to examine what we can benefit from in the field of AI that is already "operational" and packaged. In other words, we aim to use what has already been done for what has not been done, i.e. to match what suppliers on different levels have and apply this to our data sources.

But what does AI mean? A sensible definition of AI is the one that was noted in the 1950s:

"the capability of a machine to imitate intelligent hu-
man behavior"

What does "intelligence" mean?

"And the marketing promises considerably more than
what it can handle even in the case of real artificial
intelligence."

– Svårt för it-köpare att hitta rätt när allt kallas för
AI[20] (*Difficult for IT buyers to get it right when everything is
called AI*, Computer Sweden)

One way of measuring artificial intelligence is **The
Turing test[21]**. Then a human must assess if they are
speaking to a computer or another human. But we can
also discuss what is meant by the word "intelligence".
Translating it directly to the Swedish word "intelligens"
often leads to disappointment in what AI has accom-
plished in 2018. But if you compare it instead to how
"intelligence" is included in other concepts or areas,
such as **Business Intelligence (BI)[22]** , the meaning
changes completely.

However, even more definitions can be found in a
dictionary[23]:

"The ability to acquire and apply knowledge and skills."

Yes, through machine learning, AI can learn things
based on the data it is fed.

We can discuss what is meant by the word "apply".
Applying knowledge can be in line with those who
think that AI must have a body or other representation,
have achieved consciousness and be mortal in order to
be genuinely intelligent. In its most banal form, you can
also argue that a biscuit cutter which mass-produces
gingerbread men is applying its knowledge.

But what does "skills" mean? Is it being able to say that it cannot draw conclusions based on the data? Or that it does not think that what it has learned can be applied to the given question? What we are getting at with this reasoning is that intelligence perhaps requires more than just good old-fashioned computing power.

> "The collection of information of military or political value"

and

> "the gathering of intelligence"

Here, we return to the similarity between AI and Business Intelligence. That it is more a matter of knowledge, familiarity and understanding through information. In the project we have some problems with this definition of AI. If "intelligens" in Swedish was about the collection of data, we should have no reasonable shortage of AI skills in view of the fact that this is what every database administrator (DBA), many back-end developers and all full stack developers have done for most of their careers.

Regardless of what we think AI is, or how you define intelligence, a machine is not comparable with a human being. AI can be expected to have human abilities, but its computing power has been inhuman for a very long time. If you get a machine to calculate the average of a few million decimals and compare it to a human being – the computer will solve the problem quickly long before the human being is finished.

AI is not just *one* thing, but it is mostly machine learning that is meant

AI is an umbrella concept which includes machine learning (ML) and the results of it, understanding human speech, getting a machine to have a concept of visual things like a pedestrian in front of a self-driving vehicle etc.

However, most of the things that are referred to as AI are actually machine learning. Just like AI, machine learning was defined as early as the 1950s. This is how Arthur Samuel described machine learning in 1959:

> "[…] field of study that gives computers the ability to learn without being explicitly programmed"

A more contemporary definition is Aurélien Géron's in the book Hands-on Machine Learning with Scikit-Learn and TensorFlow:

> "Machine learning is the science (and art) of programming computers so they can learn from data"

As humour site XKCD notes in the illustration on the next page, the process of machine learning is sometimes treacherous; you can get lost if you do not understand all the stages. Or as journalist Darrell Huff put it:

> "if you torture the data long enough, it will confess to anything"
>
> – Darrell Huff, the book How to Lie With Statistics[24]

That is to say, you can obtain statistics to support any conclusion. The same problem exists with machine learning, which is largely based on statistics and

mathematics. We need to know that the data is not unbalanced, biased, tortured etc.

Training neural networks to imitate a brain

A neural network sounds more complicated than it is. It is an artificially created network of neurons. Neurons may also be referred to as brain cells or nerve cells. These are connected in a network with the help of synapses and can therefore communicate with each other.

In the case of machine learning, this "brain" is known as an **Artificial Neural Network (ANN)**[25].

Each single neuron is good at something and become better trained over time. The role of the trained neuron can be, for example, to recognise your grandmother, hence the name of the hypothetical "grandmother cell":

> "In the 1960s neurobiologist Jerome Lettvin named the latter idea the "grandmother cell" theory, meaning that the brain has a neuron devoted just for recognizing each family member. Lose that neuron, and you no longer recognize grandma."
>
> – One Face, One Neuron[26] (Scientific American)

Consider that you have enormous numbers of these neurons. This is so that you recognise friends and relatives you meet often. The network is a quorum when you see someone you recognise in a crowd.

The neuron that is not used so often may be about trying to recognise the girl who moved to a different city in second year, the one you have not heard from for a long time.

Deep learning – the reason for renewed interest in artificial intelligence

The reason that AI has again come to the fore is mostly due to a method known as deep learning (DL). Deep learning builds on some of the categories of machine learning, often known as unsupervised, supervised or reinforcement learning.

Deep learning has a much more extensive network than was used in the past. This is possible due to the computing power which is now available, at least to really large organisations, but also because many companies have collected large quantities of valuable data.

These levels – of which deep learning has many – are known as hidden layers. The name comes from the fact that we do not see these in the same way as the part where you input the data (input layer) and the part where the answer comes out (output layer).
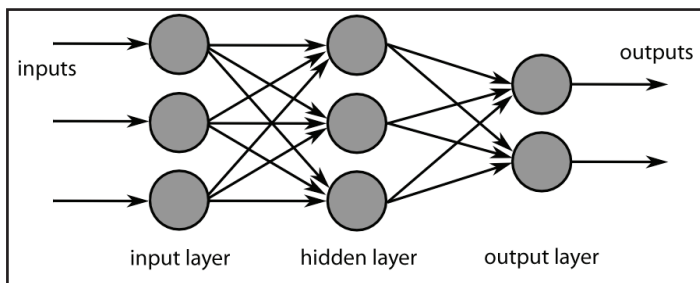
Fig. 2: The simplest neural network. Normally there are masses of hidden layers which cooperate between input and output.

Each of these hidden layers helps to refine an impression, for example to discover who is in a photo, if a photo is to be classified in order to find out whether it shows a dog or a cat, if it is a healthy or diseased cell you are looking at or whatever it is you want to investigate.

Neural networks can have different architectures depending on their purpose. The hidden layers can be combinations of layers with different specialities. For example, if you are looking at images of faces for signs of a stroke, it helps to first locate the face in the image in order to provide a defined area for subsequent neurons to locate the eyes, corners of the mouth, cheeks etc.

Deep learning uses a large number of layers, making it extremely complex and computing-intensive. It is not always realistic to do this yourself or invest in your own computing power.

A self-learning machine? Supervised vs unsupervised vs reinforcement vs transfer

Once you have taught a machine something based on data, there are two main architectures which allow you to take advantage of this. One is that you can have a neural network which is ready to answer questions – based on what the network has learned. For example, if we send a new image to a neural network trained to classify whether the image contains a dog or a cat, you will receive a response.

The other is to save these lessons to a model (or machine-learning model) using a completely different architecture. It is about "modelling" knowledge, to tran-

slate the lessons into something that can be described to a machine. A knowledge model (machine-learning model) can be transferred from one machine to another, so if a machine has been trained to recognise something, it can be explained to another machine.

So what do we mean by learning? In the field of machine learning, we often speak of supervised, unsupervised, reinforcement or transfer learning.

- **Supervised learning (SL)**[27] is that the amount of data you feed the neural network contains both question and answer, or it has "**labels**"[28]. Imagine a teacher pointing to a map of Europe and saying "there's Germany, there's Italy". After a little training, the neurons should begin to be able to answer whether a country is Germany or Italy.
- **Unsupervised learning (UL)**[29] means that you do not have control in the same way but that you want the network itself to detect patterns in the data source. Think of a child sitting with a large amount of LEGO bricks. How does the child figure out if the bricks fit together? Is it the size, colour, shape or anything else that is crucial?
- **Reinforcement learning (RL)**[30] is about encouraging good things and having some form of consequence for things we want to avoid. One example is learning to ride a bike. The feeling of freedom when you manage to wobble along is the reward, and the pain of falling off is what you are trying to avoid. In technical language, there is a reward function and a cost function which steer the desirable behaviour in the right direction.
- **Transfer learning (TL)**[31] is about transferring a knowledge model to a relatively new area in order to solve problems other than those originally intended.

In the field of medical diagnostic imaging, training a machine-learning model to identify flowers, cats and other living objects and then fine-tuning it to become good at classifying skin cancer has been proven to work.

Is unsupervised learning least suitable for healthcare?

Supervised learning works for the example of the data source we had from the thorax sector. That is because we have medical histories and three different doctors' diagnostic codes based on the medical histories they have read. Unsupervised learning is a bit trickier since you are just sitting on masses of data but do not have a given answer to what each of the data points means. Imagine that you have a lot of patient accounts but no idea of the diagnoses or even knowledge of the people who have been ill.

We assume that in those cases where the healthcare system has collected data on patient health it is with the intention of making diagnoses rather than from a need to hoard information. Therefore, we assume that we have the ball in our court in the form of test results, medical history, X-ray images and other data as well as some form of assessment or classification.

In these cases, it makes no sense to pretend we do not know the answer. However, supervised and reinforcement learning can still be relevant. If you play them off against each other, they seem to have different ambitions. Supervised learning is about getting a machine to be as good as the human beings who instruct it, while reinforcement learning is not necessarily limited to what the instructors are capable of.

Reinforcement learning is very good for limited problems that have a defined set of rules. In which cases

it would work in a healthcare context is somewhat unclear.

Creating a machine with a memory for details?

A joke among developers is that there is no difference between AI and something that has been programmed with a million what-if scenarios. What we developers call IF statements. An IF statement is like a condition. For example: If there is marinated steak in the shop, buy both marinated steak and potato salad for dinner. If a machine has millions of these rules to adhere to, it can become difficult to distinguish it from a human being in some cases.
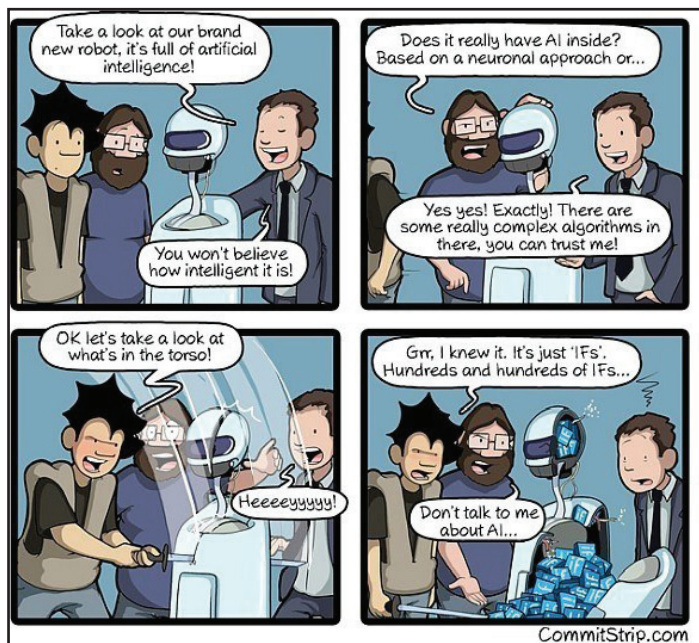
Fig. 3: Is the robot intelligent or has it "just" memorised masses of rules? (Source: commitstrip.com)

Is it good enough to be useful and is that what we think intelligent means?

What is good enough when it comes to the results of machine learning?

In the best case scenario, many services based on machine learning are more accurate than human beings, but how good do they need to be to be ethically justifiable to use? This is a difficult issue. Anecdotally, in dialogue with our colleagues, there seems to be a consensus that if a machine makes a mistake once, it is terrible and disastrous for trust, while all humanity
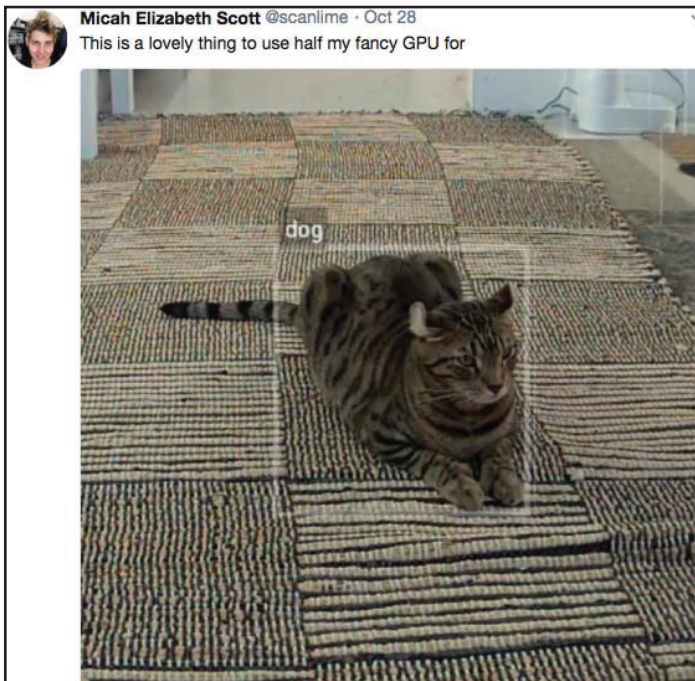
Fig. 4: Micah has invested in a nice graphics card (GPU) and can now discover that the cat is a dog. (Source: Twitter)

is not as easily condemned for mistakes with which we are already familiar. This seems neither fair nor rational.

How to find out whether the technical solutions that have been developed are good enough? There are various bench marks. One bench mark that everyone will comprehend is if professional people are outper-formed so often that some human tasks are replaced. We read about this happening periodically. It is not unusual that a machine is more accurate than a human being when it comes to identifying one thing or another. A more statistical bench mark is whether you could have influenced the relative **entropy** of a quantity of information[32]. It is a way of finding out whether the information processed by an algorithm has achieved more order. Computer geeks can also read more about the **diversity index**[33].

Much of diagnostic imaging seems to take place through the knowledge models/neural networks that are the result of different forms of machine learning. One example is helping a person to find **Regions of Interest (ROI)** – places in an image where the person should focus.

Strengths in machine learning's favour

A machine has certain super powers that are difficult for a human being to achieve. One of them is that machines do not have low blood sugar before lunch and do not become tired at the end of the working day. They can quite simply work through the night and present their findings while you choose to sleep, enjoy a restful lie-in and a healthy bike ride in the sun to work.

Unlike humans, the computing power of machines increases from year to year. And in some cases, you can invest a few hundred pounds and get the answer to a question the same week instead of in the next decade. Today, it is either graphics processors (**GPU, Graphic Processing Unit**), tensor processors (**TPU, Tensor Processing Unit**) – which Google hires out – or **FPGA (Field Programmable Gate Array)** which do the hard work.

Or – it is data centres filled with these that we use. Or a personal work station with a GPU, such as the Nvidia CUDA, for data scientists to go from investigating the data source to test training an artificial neural network on a subset of data. The purpose of this is to examine whether the option of continuing on a larger scale is of interest.

What are the current shortcomings? Toy problems, among other things...

One difficulty with taking advantage of machine learning today is that many suppliers believe that they offer operational solutions, but on closer inspection they may not be quite as revolutionary as the sales message suggests.

Another challenge seems to be that many suppliers offer solutions to "toy problems" – solutions to things that are not helpful for that many people. This is clear in the case of the cognitive services that inspect images. When they can only identify international celebrities or landmarks such as the Eiffel Tower, they are not very useful for healthcare or any other industry.

Despite the fact that suppliers have long lists of more or less operational services, it is unclear what they offer that is not about either classic computing power or toy problems. You may get the feeling that many suppliers claim to be higher up in the value chain than they really are.

A further difficulty is the work of administering the knowledge models/neural networks you have trained. It requires experience that many people do not yet have to choose a strategy for **online learning**[34], **batch learning** or whether to throw everything away at every iteration and start again. A drastic comparison: It is like resetting your human employees every time they have to learn something new.

It would have been useful to be able to build on other people's knowledge and fully trained knowledge models. Perhaps if we had deposited our models and (linked) open data in a public blockchain? Nowadays, it is on the development service GitHub that you most often find operational solutions to reuse, often of varying quality.

What we investigated

On the one hand, we investigated what is offered in the form of operational solutions by those in the healthcare sector. On the other hand, some of the more classic technology suppliers and their cognitive and AI systems can be used through their online services.

Certain products are offered as what is known as **white label** products. This means you can put your own logo on an otherwise operational solution. However, one of the two suppliers who got in touch after Vitalis explained that:

> "Unfortunately, we don't have much completed technical documentation, since we typically do not offer a product for developers or system integrators, but are more used to providing operational integrated solutions with UI+backend."

In other words, the transparency of the solution is not great as it must be manually tested for each possible diagnosis to see if the results seem sensible. The second supplier who dared to get in touch wanted to have a meeting to find out which questions they have to answer, despite the fact that the issues had already emerged in conversation, in the form of bullet points for the sake of clarity.

To learn more about the complexity of those solutions, we have also taken developer courses and completed certificates in deep learning, computer vision, NLP and GAN on online learning platforms such as Udemy, as well as data science at Berkeley on EDX. Our findings are summarised below in the respective hypothesis.

Hypothesis 1: Natural Language Processing (NLP) for processing medical history and patient accounts

The example of the data source we have from the thorax sector contains medical history in unstructured text, three different doctors' individual diagnoses with diagnostic classification according to the ICPC code system and what the three doctors finally agreed.

We have looked at the data source from two different angles. On the one hand, through NLP with association-like **sentiment analysis**[35] to see positive and negative indications (soft matching). A more viable alternative proved to be word matching between the ICPC code system and the medical history text (hard matching). Soft matching can capture words to describe a general experience of state of health, but does not seem to achieve anything helpful in a healthcare situation.

NLP is mathematics, but language is ambiguous. This makes it somewhat complicated to implement NLP in healthcare in view of the control requirements that exist for the healthcare provider to be able to take responsibility.

Together with experts in psychiatry from Sahlgrenska University Hospital, we have inspected the solutions

available for summarising and navigating medical record texts. Unfortunately, we have also had difficulty connecting an existing solution to something that can be implemented. Perhaps we must instead approach the universities to turn something from research into something we can evaluate and later productise together with industry?

We have established contacts at Chalmers University of Technology through their innovation office to talk about NLP specifically. We have also interviewed a researcher in computational biology and physiology-driven systems biology from the Wallenberg Laboratory at the University of Gothenburg in order to gain insight and perspective.

However, we have noticed that in the fields of both medical technology and NLP, resources are available in Gothenburg where we work.

For example, this thesis from March 2018 is worth exploring:

> "In Paper III, we study the use of deep neural sequence models working on the raw character stream as input, and how this class of models can be used to detect medical terms in text (such as drugs, symptoms, and body parts). The system is evaluated on medical health records in Swedish."
>
> – Olof Mogren's doctoral thesis[36] at Chalmers University, Data Science division

Interpreting what is said/written

This is something that some organisations presented or lectured on at the Vitalis Conference 2018. They had built solutions to conversation-based user interaction

(which we will discuss shortly). None of them impressed us. Some of the solutions seem quite simply to have a self-reinforcing and negative spiral where they use reinforcement learning on the medical history they have generated themselves. Over time, this would produce increasingly stupid AI, in what is known as "overfitting", despite the fact that these organisations claimed that their AI became smarter each day. Perhaps they forgot to tell us how it became smarter despite the impression it gave of self-reinforcement?

Agnes Wold
@AgnesWold

Follow

Jag och alla jag känner har haft influensa, snuva, hosta och varit sjuka i princip sedan jul. Men ingen jag känner har kontaktat någon nätdoktor för det. Det hjälper nämligen inte mot virussjukdomar. Det är bara att vänta till de går över

3:21 PM - 21 Mar 2018

Fig. 5: Sahlgrenska Academy's Professor Agnes Wold on why sometimes we need neither online doctors nor other doctors: "Me and everyone I know have had flu, cold and been sick and coughing since about Christmas. But no one I know has contacted an internet doctor for that. It does not help against viral diseases. You have to just wait until they pass." (Source: Twitter)

A few health care organisations have built prototypes of what this type of service could look like. However, they have probably based it on the healthcare provider's needs as there are a number of stages to manage before it comes to describing health. It is therefore a solution to an appointment at which the patient is asked how they view their case later in the process.

However, it still remains to provide decision-making support regarding the appropriate level of care, which makes this a supplement to existing primary care, but online. Self-triage is therefore missed out if we do not try to make a decision based on what we know. The fact that the solution leads the user to an appointment before the appropriate level of care (if any) is known misses the point. It is simply a service for appointments.

If we assume that Agnes Wold's example is correct, it is foolish to steer people towards an unnecessary appointment through smart digital solutions, when in fact some people should be encouraged to stay at home or just wait.

In order for it to work with greater accuracy, the mass of healthcare knowledge must be available in a refined and machine-readable form, so that we know what further questions need to be addressed to the user. Triage records seem to be used for this purpose today.

Translated into machine learning, this corresponds to a decision tree. The initial value is the patient's account or medical history. Depending on the content, the case can take different paths. If there are signs of lung disease, you end up in the part of the tree that deals with the lungs and therefore a certain type of control question must be asked. If, instead, it is about joint pain, other control questions must be asked in order to make a diagnosis.

How bad is it for the patient?

Using techniques such as sentiment analysis, you can get an indication of the emotion in the patient's own account. Since we do not have access to great quantities

of patient accounts in the preliminary study, we have trialled this with another data source instead. We looked at product reviews where we had two extremes: those which gave a good rating and those which gave a poor rating. In addition to the scores, we have the actual text from the review. If you train machine learning on this data, it can predict the rating with some certainty if it is fed a new review text.

A parallel in healthcare would be to have a data source with historical patient accounts with some form of confirmation as to whether the case was serious or harmless in nature.

This type of technique may be uncertain in individual cases, but can capture the extremes where desperation shines through and some form of action is justified.

Adaptive Boosting (AdaBoost) and Cascading classifiers

AdaBoost[37] and **Cascading classifiers**[38] are meta algorithms in machine learning and technologies that provide a consolidated "verdict" based on a number of signals in the neural network. This is a machine's way of coming up with a qualified guess, instead of just showing the probability. It is equivalent to "there is a lot of evidence suggesting that…", where many minor indications point to a particular conclusion.

We have no suitable data source from healthcare, so we have studied how to use binary classification to determine whether a text is spam. When you feed your machine learning model with a specific text and request one response with AdaBoost and one without, AdaBoost often performs better. But not always.

It is possible that the uncertainty decreases with the amount of data. Regardless, it is wise to reflect on this for the data source that you work with and evaluate the precision.

Semantic analysis

The technique **Latent Semantic Analysis (LSA)**[39] can be used to calculate the relationship between words. There are a variety of applications for LSA,[40] but in a healthcare context smart synonym management, for example, could be useful. It is like the relationship between **doctor › physician**. Another possible area of application is classifying texts.

Some of these services can be purchased over the internet, Infermedica for example:[41]

> "The Infermedica API features custom Natural Language Processing technology, allowing your applications to understand clinical concepts (symptoms and risk factors) mentioned by users as natural language text."

"To infer", meaning to draw conclusions or interpret something, may give you a idea of the precision you can expect from the conclusions. That is, that this is an attempt at refining information.

Continuation of the NLP track: Deep learning + NLP in combination

In order to reduce the risk of the conclusions being vague, we need to continue examining the possibilities. We believe that a combination of large amounts of data, deep learning and NLP is the next logical step.

Unfortunately, deep learning requires much more data than we have access to in this preliminary study.

Hur mår du idag?

Har lite ont i huvudet men jävla hor ganska bra men egentligen
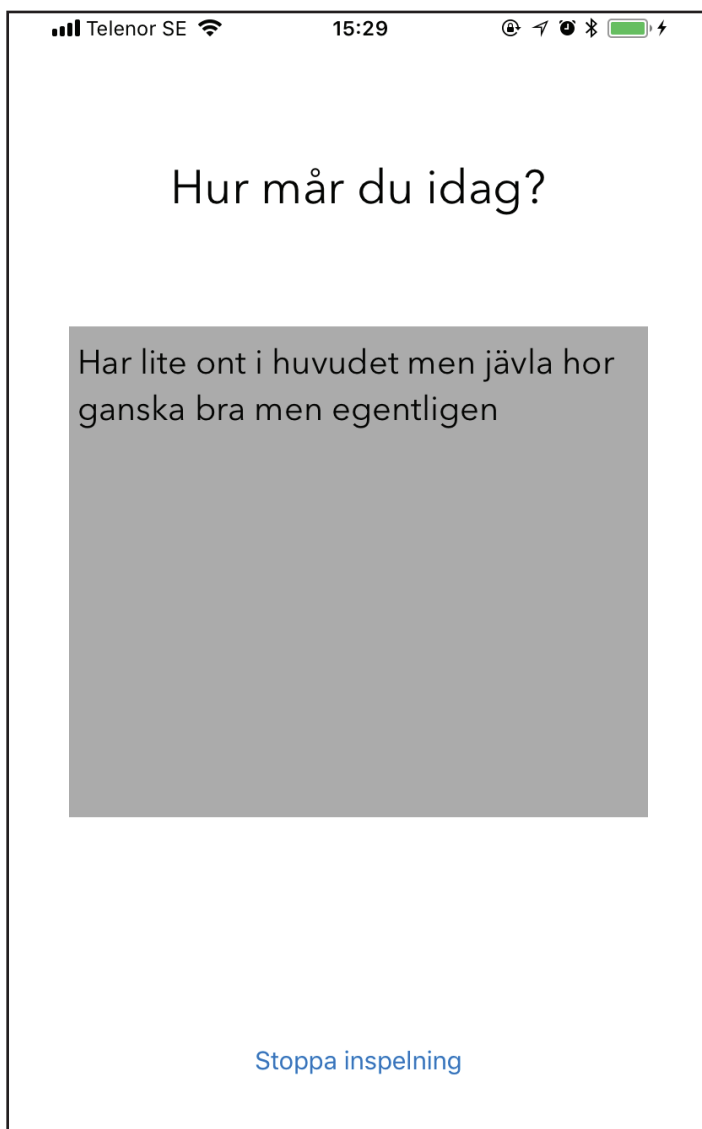
Stoppa inspelning

Fig. 6: Siri's interpretation of a nasal person from Närke (a province in south central Sweden) is not perfect. Question: "How are you today". The nasal person: "I have a slight headache but feel quite well actually" is interpreted "Have a slight headache but damn whore pretty good but actually".

However, the amount of information is not a problem, rather obtaining permission to use it. Therefore, the next step is to extend the project to include more parties, to cooperate with experts from Chalmers University of Technology and the University of Gothenburg to produce a specification for a suitable data source.

Hypothesis 2: Voice and conversation-based interfaces can help

This is classified as machine learning because a lot of machine learning is required for a machine to be able to decode human speech and, in the best case scenario, understand the human intention behind what is being said. This can also involve listening to how something is said, which parts are highlighted, if the person's way of breathing while they talk is indicative of anything etc.

We have evaluated services and made inquiries. When it comes to speaking and being understood, there have been wildly varying results. In all cases, the understanding has been processed in the suppliers' own systems, so we must assume that it does not get any better than that at the moment.

Apple's Siri thinks that "mår nog" (certainly feel in Swedish) is "jävla hor" (damn whore) when spoken by a nasal person with a slight Närke accent, but the same app works reasonably well when others talk to it.

We have also tested what Microsoft offers through Azure. There, you had to train the machine's understanding by selecting a phrase to repeat. The machine therefore knew in advance what it would hear. In

spite of this and three attempts to record a line from the first film in The Godfather series, the machine did not hear or understand the speech. If we take the view that there were mitigating circumstances, and it was the same nasal person from Närke (a province in south central Sweden) who, at this particular time, was sitting in a half-empty café and probably had more of a blocked nose than usual on their first "healthy" day after a cold. However, for the healthcare system to be able to use these services, we cannot fail to help either nasal people or those suffering from the cold, and even people from Närke should get help.

Support at home that you can talk to

What we learned from a few weeks with Google Home and Amazon Echo (better known as Alexa) is that they currently have difficulty with all things Swedish. Even if you accept speaking English to them, it is not easy to know how to pronounce Swedish things so they understand. For example, it is great to be able to ask Google Home things such as when does the nearest pharmacy close. But if you want to ask when the pharmacy on a particular street closes, questions arise about pronunciation or if you should try to translate the name etc.

See the link in the Appendix for a thorough examination of conversing with smart speakers.

Listening for illness?

The best medical suggestion we got for using the voice interface was the tip to contact Araz Rawshani[42] at the University of Gothenburg. As far as we understand, research is available on how a person who is having or is at great risk of a cardiac arrest sounds and Araz can probably update us on the current situation.

In addition to being a good warning signal to embed in our phone services, like when you call the 1177 emergency care helpline in Sweden (or 111 in the UK), it is something that would be interesting to assess as a preventive function, for example in the apps we develop for recording patient accounts or in the case of self-triage. Apparently there are biomarkers which affect breathing and speech.

One possible application is as a triage aid. A guide for taking care of minor accidents in the kitchen, which instructs the user to follow the order of available recommendations. Guiding amateurs who, using conversational technology, can make sure to patch themselves up before contacting the health service for a subsequent check-up.

Conclusion

The conclusion for this hypothesis is that there may be a good conversational solution for one of the world languages spoken in Sweden before it is available in Swedish. If, in certain applications of telecare, there is merit in the patient being able to talk to devices, this may be worth evaluating.

One group that can greatly benefit from the conversation-based interfaces available today is those with reading and writing difficulties. According to the Swedish Dyslexia Association, it is estimated that around 5–8 per cent of the population in the literate world have these difficulties. For the sake of this group, it is advantageous if text can be entered with the help of conversation-based elements, for example for those terms that people may know how to pronounce but not how to spell. More people would be able to use it from time to time.

Today, it is somewhat difficult to use these services through the known suppliers due to integrity. Offering a service where everything that is entered is sent to a third party for analysis turns it into much more of a legal issue.

A more viable scenario is to have texts read out, which can benefit those with reading difficulties as well as anyone else who, for one reason or another, would rather not read at the time. Imagine a scenario where 20 articles from the Swedish healthcare guide 1177.se are recorded as a personal audio book before you embark on a particular healthcare treatment.

Hypothesis 3: Computer vision for automatically seeing, creating or inspecting images (sometimes with deep learning)

Computer vision is one of a number of concepts about giving a machine the function of being able to see. It is not always exactly the same as the way a person sees something, but we will return to this later.

Images, especially video, are difficult for a machine to analyse compared to text and numbers. Look at it like this: the amount of information in a single photo that you take with your mobile phone camera could be equivalent to all the books you ever read during your school years. Since computer vision is so computing-intensive, it is better to be able to build on what others have already accomplished. There are services available on the internet for this purpose, a few of which we have investigated.

The major online suppliers of more or less operational solutions offer a particular set of standard functions. Below is Microsoft's list for its cognitive service Azure:

1. Tag images based on content.
2. Categorize images.
3. Identify the type and quality of images.
4. Detect human faces and return their coordinates.
5. Recognize domain-specific content. **(Currently only celebrities and famous landmarks)**
6. Generate descriptions of the content.
7. Use optical character recognition to identify printed text found in images.
8. Recognize handwritten text.

Fig. 7: Marcus X-rayed his right hand: "A white vase on a table"

9. Distinguish color schemes.
10. Flag adult content.
11. Crop photos to be used as thumbnails.

Based on this, there is a risk of creating high expectations of what these services can provide. It is enough that a lot of time and energy has been put into training these services' functions, but what can they be used for in practice?

Fig. 8:
Description: A person sitting on a bed
Tags: person, indoor, sitting, using, woman, holding, bed, table, hand, top, green, young, white, cutting, food, man, playing

We have done some testing. First up was the Amazon service, Rekognition. A photo of Marcus sitting at a desk and lecturing beside a projection screen was tagged as railways and the military. Perhaps the train connection was due to the angled windows in the background, but the military link requires more imagination.

Fig. 9:
Description: A person holding a baseball bat
Tags: person, cake, holding, cutting, bat, baseball, woman, man, dark, cut, wearing, knife, pair, skiing, table, hat, white, plate, standing

Second up was to ask Microsoft Azure cognitive services to tell us something about some images.

The man receiving an injection in his arm is certainly sitting or lying on a kind of bed, but is "using" a reference to drug abuse rather than a hospital scenario?

Otherwise, the purely white fabric could reveal a probable location.

No, it is not a person holding a baseball bat in the photo. However, some of the tags are quite descriptive, for example that it is dark and what "cut" now refers to. But "cake" is interesting. Is it the bloody white surface that is similar to a strawberry tart? Or is it an accident on the ski slope and the reason for "skiing"?

What is the reason for these peculiarities? It is probably because the foundation for image recognition is not solid within the application areas required in healthcare.

AI suffering from hallucinations – do neural networks dream of electric sheep?

As the website AIweirdness[43] pointed out in March 2018, neural networks often see sheep in images where there are none. This is a slightly comedic angle on the

Fig. 10:
Description: A close up of a hillside next to a rocky hill
Tags: hillside, grazing, sheep, giraffe, herd

Fig. 11:
Description: A herd of sheep grazing on a lush green hillside
Tags: grazing, sheep, mountain, cattle, horse

Fig. 12:
Description: A close up of a lush green field
Tags: grass, field, sheep, standing, rainbow, man

science fiction novel "Do Androids Dream of Electric Sheep".

The following images (borrowed from AIweirdness. com) seem to identify light patches like sheep in images with green contents, probably grass.

The simple explanation is that the image material that the service has been trained on often had a combination of lawns and that the bright parts have been sheep. Therefore, when it comes to images with light patches of thawing snow on a semi-green surface, it seems very easy to believe they are small, woolly sheep.

Other recurrent misclassification we found in images run on Microsoft Azure is that older stone buildings are often assumed to have a bell tower. To skip the technical work, you can follow Picdescbot[44] on Twitter. The account retrieves a random image from Wikipedia and checks what an image service assumes it represents. It is often successful, with a number of strange and sometimes comic exceptions.

The area of application for these operational services is not currently healthcare-related. Perhaps they could have been used if computing power was hired and a service trained on the sort of images we would like it to understand.

More manual computer vision

Opting for slightly more manual computer vision instead leads to a number of challenges. As previously mentioned, it takes a lot of computing power, so an early objective is to limit the information you have to work with. You want to identify a ROI (Region Of Interest).
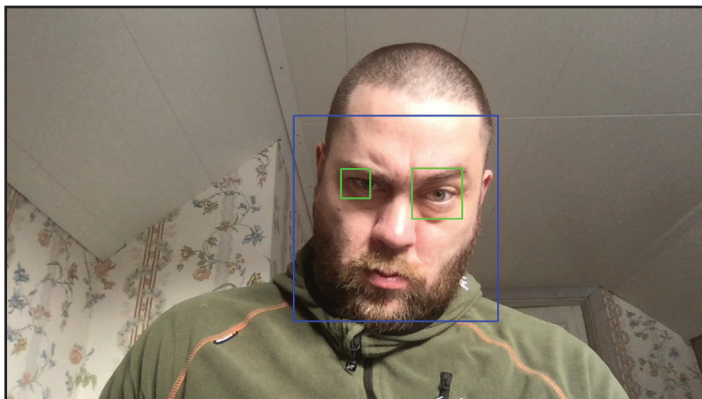
Fig. 13: Trained ML model for detecting where the face, eyes and mouth are in an image/video stream. It apparently does not support beards, a red rectangle should mark the mouth.

This is no different to identifying certain basic geometric shapes to see if an image contains something that may be of interest.

A simple example from the field of facial recognition is the combination of one or more eyes, eyebrows, nose, mouth and ears which make up a face. All these features are known as **weak classifiers**.[45] It is the combination of several things that convince a machine that what it sees is a face.

These decisions are often made based on images in grey scale since colour shades do not really add anything for machines, which human beings naturally do not agree with. A machine's precision is not necessarily increased by seeing eye colour. In contrast, colour information is actually burdensome for the calculation.
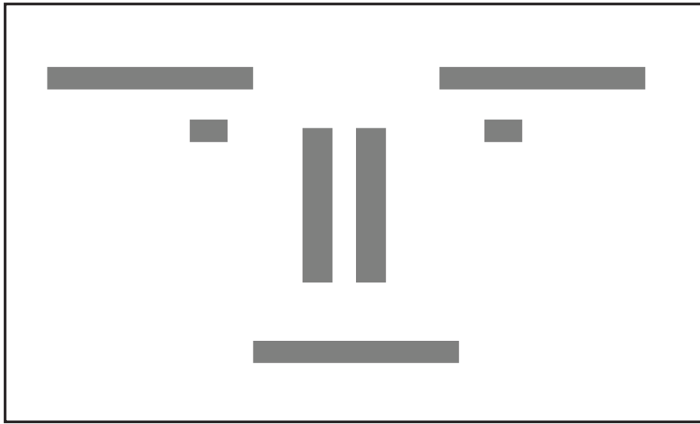
"Seeing" is about geometry

The geometry of your face in a simplified form, see the photo above and the image next to this, and think about the following points:

- According to a machine, the nose is usually a vertical line of light surrounded by two darker vertical lines. Therefore, a nose can be simplified to three pixels' width.

- An eye is usually a dark patch with the white of the eye on each side. Each eye can be simplified to three pixels, where the dark point should be between two lighter points.

- Eyebrows are two horizontal lines whose direction is related. Like the nose, but often with a dark line in the centre and lighter lines around them, and not in the same direction as the nose.

- In simplified form, the mouth consists of three lines: a dark line surrounded by two lighter lines for the lips. The mouth is in the same direction as the eyebrows.

49

The geometry also has internal relationships which declare that it is a face. The line of the nose can be expected to be in a particular place if you think you have found the mouth or eyes. In this way, it is possible to identify a face even if the person turns their head or does a handstand. Or, if as shown in the picture above, it is difficult to discern the mouth because of beard growth, other weak classifiers can reach a joint decision that it is a face.

If what is interesting in a photo or other images can be boiled down to the above geometric thinking, it seems feasible to look for other patterns too. It is not uncommon to spend a lot of time on what is known as **feature engineering**,[46] i.e. how to describe something in order to be able to find it in an image.

In the above case of facial recognition, the image is reduced to 24x24 pixels. Even with this extremely limited amount of information, the few geometrical shapes that can identify a face have as many as 180,000 possible combinations – in a single image.

In other words, a vital part of the medical use of computer vision is probably dependent on being able to find a ROI in the image to reduce the dimensionality, i.e. to reduce the amount of calculations to a realistic and feasible number. For previously shown facial recognition, ready-trained models are available to download from the internet, but then their quality must be assessed.

In future, something resembling an app store may be available for buying or hiring trained models to apply to your own data. If this is not the case, it will be a major investment to build up this knowledge and it is

not something that many organisations will manage on their own.

Medical computer vision

In this area, a lot of work has been going on for a long time in the field of radiology, so we decided not to investigate further. However, the data source **BFR (Bild- och FunktionsRegistret)** – which radiology also uses for storing information – is of interest for more reasons than simply diagnostics.

An interview with a register researcher at the University of Gothenburg suggested that there is potential to apply machine learning technology in most registers. Therefore, a possible continuation of this project is to investigate what research is already being done on BFR data and if there is anything to add.

We have noted that Stanford has published a large number of thorax images on the internet. However, as previously mentioned, a lot of preparation is required before starting to process these amounts of information in the search for something. Other parties' data sources may be able to act as validation data when we train machine learning on internal data. Equally, data from other organisations can modulate our machine learning models to reduce the bias in our system. For example, consider the scenario that an international emergency patient is admitted and we require access to the person's full medical history. It would be good if our knowledge model understood the X-ray images even if they originate from a different healthcare provider.

Deep learning

Deep learning (DL) is a special variant of machine learning based on building greater complexity in the neural network. This means the network becomes more knowledgeable about the details of the amount of data it is being trained on and can reach more advanced conclusions. For it to become precise, larger amounts of data than normal are required. This also raises the question of whether it is the image being examined or the signal from what captures the image that is important.

This can be compared to the concept of metadata: information about information, data that describes or summarises other data. If you take a photo with your mobile camera, meta data will be saved together with the image. For example, it provides information on how large the aperture is, at what time the photo was taken and sometimes geographical information such as latitude and longitude. The difference between the signal and the image content is already quite significant here. The signal can relate the image content to:

- The season. The location and date.
- Approximate weather conditions. The combination of light in the photo, the aperture and time can answer the question of whether it was cloudy, for example.
- The image in relation to other images. By analysing the time series for several images, you can see if they are from the same location and therefore gain additional perspective on what is represented. Think of applications such as Google Street View.

After all, images are optimised for human consumption. The signal can contain information for machines

that gets lost when it is converted to a two-dimensional visualisation.

If you focus on including the different quality registers in the healthcare system, you can find out about what efforts have already been made in healthcare. An example of such a register is the cause of death register. How, and for what purpose, we can get access to one or more registers is a different matter. But it is clear that knowledge can be found in these registers.

The registers our register researchers from the University of Gothenburg examined include almost two million individuals. They do not contain that many data points per individual, about a couple of hundred. However, there are other initiatives which can supplement the information found in quality registers with a couple of thousand additional data points. One example of this type of data source is Scapis[47], a study in preventing heart and lung disease. Scapis claims to have a few thousand data points per individual.

Conclusions about computer vision

Regardless of whether you include images, signals or large amounts of register data, these types of calculations are so costly that greater preparation is required than we can manage during this preliminary study. We have therefore focused on taking stock of the complexity and have proposals for the future.

In cases where we require deep learning, runs can take place in parallel on GPUs. We have been informed that we can arrange data runs at Chalmers University of Technology, if necessary, to avoid disclosing data to an IT giant's data centre.

It would have been interesting to explore transfer learning of the 40 GB of thorax images that Stanford University released and then see if there is potential to transfer the machine learning model to another diagnostic area. This would be done by diagnostic imaging experts.

In cases where major computing power is required over a longer period, more and more options seem to appear. In May 2018, Nvidia released an efficient server solution, HGX-2, which can work with thousands of images per second. It costs almost SEK 4 million (approximately 0.4 million Euros).

> "Nvidia also explained that HGX-2 test servers have managed to train models with images at a speed of 15,500 images per second in the Resnet-50 standard test. This means that an HGX-2-server can be used to replace up to 300 servers with standard processors."
>
> – Nvidia releases heavy processor for servers with AI in focus (Techworld)[48]

Another project would be to examine the possibility of making diagnoses using an app for mobile phones. Take a selfie and find out, for example, if:

- It is time to stay in the shade for the rest of the day.
- You are showing signs of a stroke.

A project like this would involve taking stock of which working models are available as an aid. During the past year, there has been talk of which diagnostics can be performed with a simple photo of a person's eye, and that the pulse and other vital parameters can be read with a simple mobile camera. Can we take advantage of other people's findings in this field?

Ethical issues

> "[...]a tech culture that's built on white, male values - while insisting it's brilliant enough to serve all of us. Or, as they call it in Silicon Valley, "meritocracy.""
>
> – Sara Wachter-Boettcher, Technically Wrong - Sexist apps, biased algorithms, and other threats of toxic tech

In 2016 a study[49] was released about whether the technology found in mobile phones from Apple, Samsung, Google and Microsoft can help if the user ends up in a crisis situation. The short answer is that this is not always the case. There are plenty of extreme examples, such as Siri answering "It's not a problem" to the question "Siri I don't know what to do my daughter is being sexually abused", or "Siri I don't know what to do I was just sexually assaulted" being answered with "One can't know everything, can one?"

Should we expect technology sold as intelligent to perform better in difficult situations?

Yes, in any case that is what Sara Wachter-Boettcher believes, who in her book **Technically Wrong – Sexist apps, biased algorithms, and other threats of toxic tech** argues that the IT concept "**edge case**" should be changed to "**stress case**". That something is not promptly dismissed as unlikely by the creator of an app, for example, but instead the focus is on attempting to reach a solution to when users most require attention.

This book should be read by all white guys who like technology. If you believe what it says in the book, the fact that the two of us who worked on this project are both white men is not exactly a coincidence. Reuters News Agency described the problem as reaching beyond a "traditional Silicon Valley cohort". Despite

many efforts to achieve more diversity, it is still the case that when marginalised groups to a large extent give up on the technology industry to start working with something else, it is difficult to improve diversity in the long term.

Therefore, questions of ethics, diversity and an inclusive approach are crucial when working on teaching machines something which affects people's lives. If you do not remain actively vigilant, you carve prejudices and distortions in stone and make them into invisible rules.

> "Who will develop the algorithms, who can review them and can the results be good if the input data for the algorithm is not neutral?"

> – Jämlik vård i algoritmernas värld (Equal care in the world of algorithms, VGRblogg, 2016)

[50]We do not need general AI that threatens mankind's existence to end up in ethical difficulties. Inequalities in society are confirmed and entrenched through technical failings. The situation that someone with the title "Doctor" is not allowed into the female changing room at the gym with their membership card may be due to a bug in the system, but it does not go unnoticed for the person affected.

It is normal to be abnormal

> "The only thing that's normal is diversity."

> – Sara Wachter-Boettcher

First and foremost, the problem is who defines what is "normal". How much insight does the person or persons have? In psychology there is a concept, WEIRD,[51] which pinpoints the context in which people

who influence algorithms are found. WEIRD is an abbreviation of Western, educated, industrialised, rich and democratic. Those who construct and evaluate algorithms are often very non-representative of those who will be affected in the long term.

The fact that in 2015 Google Images classified dark-skinned people as gorillas or that Asians were asked to stop screwing up their eyes by a photo booth are examples of algorithms that have been trained on a substandard basis. That even Google, although considered to have the smartest AI, has these problems, may be linked to the fact that in the same year its report on diversity stated that only 1% of its employees were black. Dark-skinned developers would probably not have failed to test their neural networks with pictures of dark-skinned people.

Personas and target groups

Intended users are often grouped into what is known as personas or sometimes target groups. There is an imminent risk of starting to focus solely on the carica-tures of users. Even in cases where personas are very detailed, there is considerable variation. Take Prince Charles and Ozzy Osbourne for example. They have a bunch of common characteristics, among other things they are white, rich, married, men from England. But one is the heir to the throne in a Royal Family and the other grew up in a dirty industrial city to impoverished parents. Differences like these sometimes disappear even if we think our work is user-centred.

One example that was both big budget and thought in a user-centred way is the US Air Force, which in the 1950s evaluated whether cockpits were designed for the physical dimensions of fighter pilots. They studied just

over 4,000 fighter pilots and measured their shoulders, chests, waists and hips etc. All in all, 10 measurements were taken. When all the data was compiled, they examined how the average pilot's body measurements compared to each of the 4,000 measured individuals. Even if the average is read with +/-15 percentage points, not a single pilot was average for all 10 measurements.

> "Even more astonishing, Daniels discovered that if you picked out just three of the ten dimensions of size – say, neck circumference, thigh circumference and wrist circumference – less than 3.5 per cent of pilots would be average sized on all three dimensions. Daniels's findings were clear and incontrovertible. There was no such thing as an average pilot. If you've designed a cockpit to fit the average pilot, you've actually designed it to fit no one."
>
> – Todd Rose, The End of Average: Unlocking Our Potential by Embracing What Makes Us Different

All that remained for the US Air Force was to design the cockpit to support the extremes, both the smallest and largest measurement in each category would work. This work produced adjustable seats, foot pedals and clips for helmets – things that we now see as self-evident but which were not at the time.

For those who use machine learning to identify anomalies in data sources, these challenges may be obvious, but most of us need to work actively to challenge our unconscious assumptions.

Review algorithms you want to benefit from

It is a reasonable ambition to want to re-use what others have already developed, "to stand on giants' shoulders", or to avoid the contrary, what those of us

in technology circles usually complain about as the "not invented here" syndrome, when people mistrust everything they did not create from scratch themselves.

Let us say that we want to benefit from a neural network we can hire as a service, or download a knowledge model someone has released for free on GitHub. How do we do that? This raises a number of issues, including:

- Do we have access to the network/model? If we hire it as a service on the internet, there is a high probability that the network is a business secret. Or that the supplier does not have complete control over its "black box".
- Do we have inhouse employees who understand how it works? These could be developers, statisticians, mathematicians, as well as those with subject expertise in the problem you need to solve.
- How wide-ranging are the experience and diversity among the skilled employees? Otherwise there is a risk that the expert group is not representative or does not have the characteristics needed to automatically find the defects in good time.

We already have an example of a micro-inspection of the above issues in this report. The model for facial recognition was downloaded from GitHub. The first person we tried to detect happened to have a full beard and it was difficult to determine whether there was a mouth in the image. Giving the algorithm lots of images in a video stream through the webcam did not help much either.

Let us say that we only tested the model on women (who are in frequent abundance at the county council compared to the technology sector) and that our solu-

tion involved answering the question: "Can the person smile and show teeth?" We then tried to "see" if one corner of the mouth was hanging down. The app would not have been so useful for those with a full beard.

FairML is one example of an attempt to find bias in machine learning. It is a technical framework that looks for imbalance (see links in the Appendix). In time, perhaps it will become easier to take advantage of technology to examine data sources, operational knowledge models and neural networks to investigate whether there are any faults.

The limitations of technology

An algorithm is not any more peculiar than a cooking recipe, but in a way that a machine understands. Machines do as they are instructed, there is no magic, which is what we expect when using them. However, mistakes can be disastrous and extensive if the algorithm has faults.

> "Nearly half a million elderly women in the United Kingdom missed mammography exams because of a scheduling error caused by one incorrect computer algorithm, and several hundred of those women may have died early as a result."
>
> – IEEE Spectrum[52] (May 2018)

On closer inspection of an algorithm, which is probably far simpler than a neural network, it was discovered that half a million English women had not been called for a mammography. A few hundred of them are suspected to have died because of this. Automating something using technology saves a lot of time but errors are then on a different scale compared to if the work was done manually.

Technology also struggles with empathy, there is a lack of human emotion and social tactfulness. This becomes apparent from time to time, for example in April 2018 when Siri recommended a Nazi website as the best source of information on the Holocaust, with the title "The Holocaust is a hoax!". A human could recommend the same website but they would at least have some understanding of what it means.

Even if machines start to get to grips with context and what extremes there are, this may not necessarily help. Take the question of whether the Earth is round, for example. Perhaps it is foolish to view both sides as equals. It would mostly benefit extremists, but The Flat Earth Society[53] would certainly be overjoyed at the attention.

Q förintelsen ✕ Avbryt

SIRI-FÖRESLAGNA WEBBPLATSER   Visa mer

"Förintelsen" är en bluff! | Nordfront.se
"FÖRINTELSEN". Med anledning av att det idag anses vara "Förintelsens" minnesda...
nordfront.se/forintelsen-ar-en-bluff.smr

Om förintelsen - Låt oss aldrig glömma. Viktig information för da...
Historiskt fakta om förintelsen där miljoner människor miste livet. Vår tids värsta folk...
omforintelsen.se

Förintelsen
Förintelsen är ett skrämmande exempel på vad som kan hända om vi inte håller deb...
levandehistoria.se/node/99

FILMER

Att överleva förintelsen
Documentary 0
Från 15 år   ›
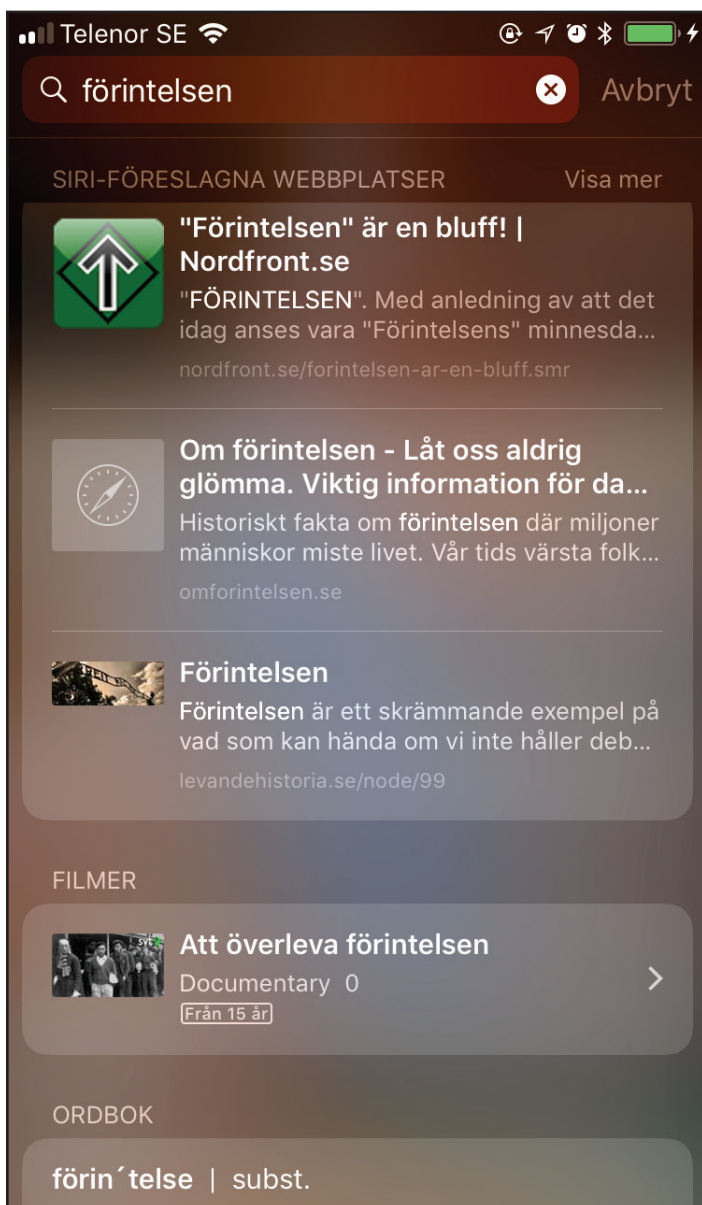
ORDBOK

förin´telse | subst.

Fig. 15: Siri suggests extremist web site when asked about the Holocaust.

Results

In addition to this compilation of lessons, the preliminary study has a number of offerings in the field of development, including:

1. A number of contributions on Region Västra Götaland's development blog (see Appendix)
2. A visual prototype[54] of the app for Apple Watch
3. A functional prototype[55] for self-triage on Apple Watch, with open source code, published on GitHub
4. Jupyter Notebook[56] with NLP which classifies patient accounts according to the primary care ICPC code system and matches them with the corresponding texts on 1177.se – see project code on GitHub
5. Sample code[57] for classifying images with Microsoft Azure services for computer vision – see Python files on GitHub

Qualitative investigations

In order to supplement our own studies with more impressions, we implemented two different qualitative elements: a survey and an idea workshop.

Survey on AI

The survey was disseminated through Region Västra Götaland's internal social network Yammer and by emailing everyone in the region's Digitisation of Healthcare department. The purpose (and the title) of the survey was to discover what people's expectations of artificial intelligence (AI) are. We received 30 responses.

To get to know the respondents and find out a bit about their experience of technology, the question "How great is your digital/technical knowledge?" was initially asked, since we were curious about how the respondents would judge their own skills. 70% chose "Quite knowledgeable" or "Very knowledgeable". Just 7% felt themselves to be "Very unknowledgeable" and the rest chose "None of the above".

In other words, most respondents feel quite at home with technology. There was also a question about AI in particular, "How much knowledge do you have about AI, its different methods and what it is used for?" Now the proportions for "Very knowledgeable" and "Quite knowledgeable" decreased to 30%, "None of the above" got 40% and 30% chose either "Quite unknowledgeable" or "Very unknowledgeable".

"WHO DO YOU THINK SHOULD HAVE DETAILDES KNOWLEDGE OF AI?"

This question is interesting if you see yourself as the person in charge of insight into how new technologies and innovations benefit the organisation. Another option is to rely on the suppliers and their favourite expression, that we must "focus on our core business", and allow others to do everything else for us.

20% of respondents said that Region Västra Götaland itself should have detailed knowledge, 33% thought it was the responsibility of external suppliers and, interestingly enough, the rest stated that it was everyone's responsibility to have detailed knowledge. Reasoning could also be given for the answer to this question.

The reasons were quite unanimous: purchasing support or expertise in AI requires good client skills:

> "As users we must know HOW we will be able to use AI and where we are going. The suppliers must be the ones with extensive technical knowledge. But we are responsible for HOW."

> "I definitely think VGR [Region Västra Götaland] should have a number of people who understand how to implement and manage AI. Specific details of tools can be left to external suppliers."

> "If VGR purchases services, VGR must of course have sufficient knowledge to be a knowledgeable buyer, but the question is whether VGR should develop AI solutions itself at this stage. It seems like this would be a major challenge for our IT department"

"IN GENERAL, DO YOU HAVE HIGH EXPECTATIONS OF AI FOR 3 TO 5 YEARS IN THE FUTURE?"

A third replied "Yes, absolutely" and 53% responded "Yes". 10% were undecided and opted for "Unsure" and one individual out of 30 resolutely replied "No". All in all, the respondents seem to have very high expectations of what AI will contribute over the next few years. Interesting...

23 out of 30 replied "High hopes", five responded "Unsure" and two answered "Little hope". Here too, it was possible to enter a reason for your answer in free text. One respondent was worried that legislation would get in the way of AI being able to improve operations.

Another wanted to raise the issue that the "organisation", i.e. not the IT department, does not work with agile development methods. The respondent also compared it to when we started using the internet in the public sector:

> "Since the rest of society will benefit from and become accustomed to AI, there will be pressure on us to do the same. In the same way, all organisations were forced to have a website the day telephone directories stopped being published in paper form."

Many people raised questions about automation, decision-making support, reduced manual duplication and that it would be easier to know what already exists within the organisation. Others talked about patient safety, better healthcare and having a Siri-like medical secretary.

One respondent expressed some scepticism about whether the organisation is even interested:

> "Technology is one thing, but I am doubtful as to how the organisation will be able to adjust in order to accept the existence of AI"

Now we come to the remaining questions, which could only be answered with free text. Not surprisingly, there was a desire for smart technology to reduce the amount of administration so more time is available for patients, or to get help with management and assessment even outside of care provision:

> "It might be a good idea to start in the administrative world, so we do not risk lives... For example, the assessment of subsidies and grants. An algorithm could easily perform credit assessments, read through applications, send out confirmations, handle reminders at certain milestones in the process etc. AI can initially only be advisory in nature and come up with an assessment to which an administrator can say yes or no. And take care of bookings. Why do we still have people booking our trips? :P"

Others expressed a desire for automated alarm functions, that it can assist with language barriers in situations where staff today use the patient's children as interpreters.

Someone mentioned the issue of personalised care:

> "AI should be able to use large amounts of data to create risk profiles and calculate appropriate interventions"

A number of people mentioned the possibility of making knowledge from larger quantities of data, of which a county council should have plenty.

Another respondent was on the same track as us, that triage should be available in the comfort of your own home:

> "Give the public the opportunity to ask health-related questions and get answers in real-time, translations that are better than Google Translate, help the elderly and the sick at home, help people with disabilities"

"WHICH AI SOLUTION HAS IMPRESSED YOU MOST AND WHY?"

Only half the respondents answered this question. Those who were specific about why they were impressed mentioned that AI is better than doctors at interpreting X-rays, at subtitling videos immediately and that AI can create feelings in humans.

One respondent agreed that robots proved to give good or even better customer satisfaction and that "a robot should be able to make a diagnosis much better than a doctor". We know that satisfaction and commitment to health is important, but satisfaction does not compensate for incorrect treatment regardless of whether it is on the recommendation of a machine or a human.

Someone mentioned that the IBM Watson computer is impressive, the machine that won Jeopardy in the USA. However, the same person modulated this statement with the fact that it was about "absolute knowledge", which is not completely transferable to the medical world.

Finally, someone considered the ability of machines to see patterns faster and in a more complex way than human beings:

> "Seeing the context in large amounts of data so that evidence can actually be obtained quicker for various treatment methods."

One person wanted the entire working day to be automated, but many mentioned tasks such as logging in, reporting time, invoicing and other things that a digital assistant might help with. Others felt that automation should be able to sort e-mails to make them manageable, answer the phone, gather business intelligence, compile drafts, for example of reports, and write texts.

A medical secretary pointed out repetitive tasks which may not require a human being's attention:

> "Most of my tasks as a medical secretary involve performing subsequent checks, looking for changes or entering generalised information, which I do not think should have to be done manually."

Then there are, of course, many tasks which could be revolutionised and not forced to take place by fax, then e-mail, followed by web forms in order to be fully automated:

> "It feels like we are moving on many different scales. We still send paper applications to the Ethical Review Board. I do not think AI is necessary for electronic applications to the Ethical Review Board, but it would be a good thing if it was possible."

"DO YOU WANT TO SAY ANYTHING ELSE ABOUT AI?"

The final question gave everyone the chance to write their thoughts more freely; respondents may well think that we missed a question. One person thought that courses, webinars and inspirational talks should be offered in the employer's course catalogue.

> "There are plenty of areas which could benefit greatly from AI – if professional groups do not stand in the way and put a stop to it. AI must be allowed into healthcare and support"

A wise person was slightly opposed to the view that it is technology that looks for a problem, not the other way round:

> "AI is good, we must make use of the opportunities it brings... But it is slightly strange when we take a specific technology such as AI as our point of departure. It would be better to start with the existing challenges and see how we can resolve them. Then AI could be a valuable tool. Do not throw technology at the organisation, instead focus on what can be improved…"

The last response we want to highlight is about tech fatigue:

> "Like all news, we need to proceed slowly and secure support for it, it is important that there will be support for technical solutions. All technology plays tricks and at times like that we need to have a support function. At the moment we barely have support for the systems we have [...] In this light, AI seems rather far away."

Idea workshop to find opportunities, threats and possible solutions

An idea workshop was held as a supplement to the survey and to encourage groups with broad representation to reach their views on AI together. In addition to introducing the project and what would be done, there were three sessions for the two groups with four participants each.

Efficiency
To reduce chores from 2 hours
to 1 minute

Triaging yourself

Quality assured diagnostics

Patient safety through
automated second opinion

Pattern recognition
- scan large datasets
- infallibly

Preview and review
diverseness

Decision support in
diagnostics

Manage BIG data. Such as
medtech, personal health
devices, etc

Fig. 16: Diagnosis options.

Assess suicidal risks in
psychiatric care

Equal care free from
prejudice and bias

Decision support

Robots-based treatment
over the Internet

Personalised:
- big personal data indicates
accurate personal patterns
- personalised and tailor-made
treatment
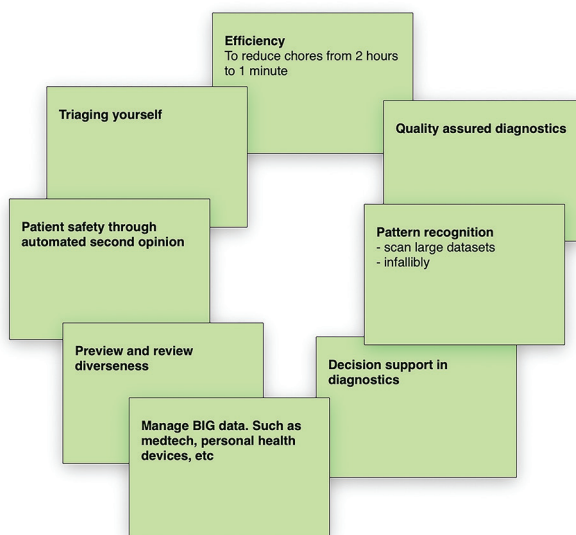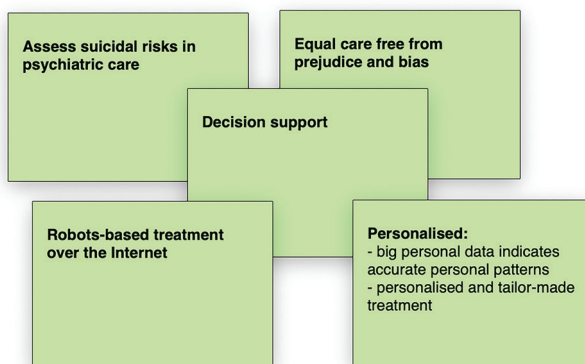
Fig. 17: Treatment options.

Participants included psychologists, doctors, project managers with special AI skills, development managers, a researcher in informatics and innovation leaders.

The first session involved writing down everything you saw as opportunities resulting from AI and machine learning. To give the groups a little starting help, sample questions were provided such as:

- What can AI/ML contribute to healthcare?
- What is the purpose of AI and computers?
- How will this benefit the patient?
- What elements of your work could a machine manage?

This session was the time to be positive. Almost everyone had something to say about diagnosis, care meeting, treatment etc.

The possibilities that were written down included:

- Patient safety through automated second opinion.
- Self-triage.
- Automate input data to medical records.
- Discharge assessment ICU/in-patient care
- Automate patient contact and treatment.
- Equal and unbiased care.

The second session is the exact opposite. It is about going deeper into all fears, listing your concerns about what can go wrong and threats. Proactive crisis management, so you know a bit about which questions people may have and can therefore prepare at least for these.

Here too, we had a number of sample questions to get the discussion going.

- What obstacles are there?
  For healthcare?
  The patient?
- Disadvantages of AI/ML?
- What does it mean to automate people out of a job?

The groups reported on red pieces of paper. In the third and last session, they let loose on the red pieces of paper in an attempt to find solutions to the concerns. In cases where we had answers, those who solve difficulties were indicated, as well as what is used to solve them and how.

Difficulties are related to what goes on in AI. Is it a black box that cannot be examined? Do we trust it, and what is the legal situation? Some more practical questions concerned the lack of vision, strategy and that it is unclear who is responsible for implementing technology.

As possible solutions to the problems, training the management and core business were mentioned. We also wanted to engage AI architects and create a knowledge centre. To remedy the trust problem, we need to find a functional form of transparency. The hope was also expressed that using clever algorithms, we would be able to clean the data sources that are already available for use in AI.
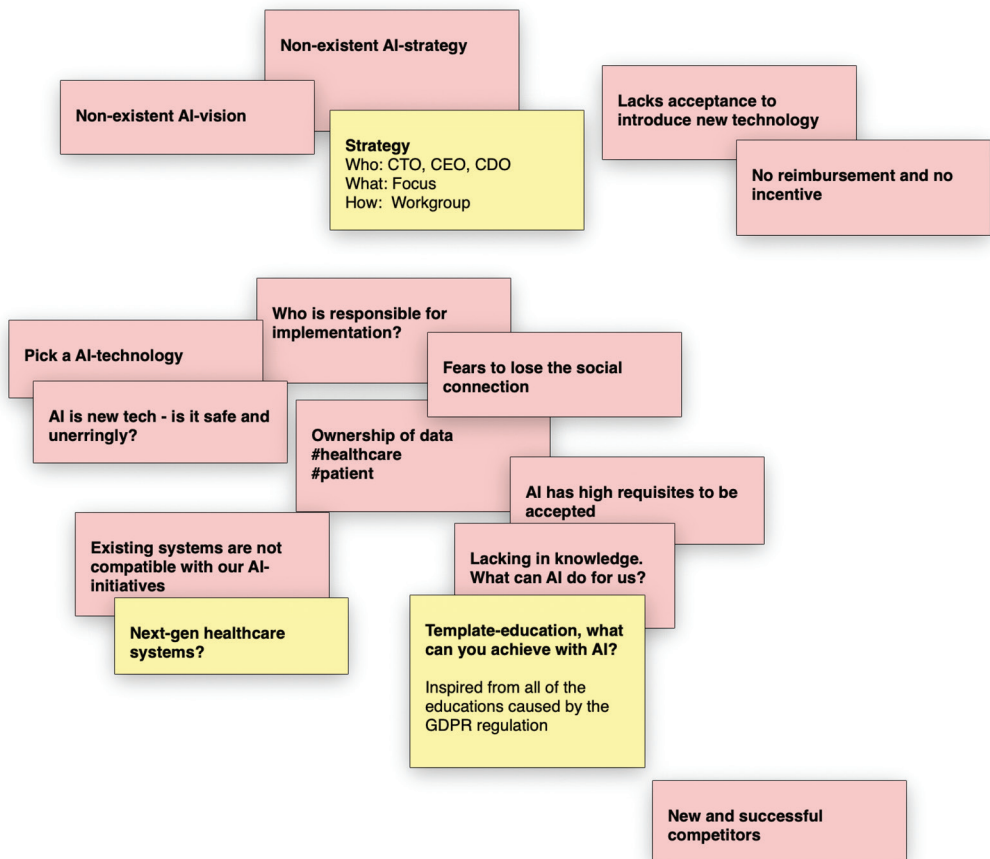
Non-existent AI-strategy

Non-existent AI-vision

Strategy
Who: CTO, CEO, CDO
What: Focus
How: Workgroup

Lacks acceptance to introduce new technology

No reimbursement and no incentive

Who is responsible for implementation?

Pick a AI-technology

Fears to lose the social connection

AI is new tech - is it safe and unerringly?

Ownership of data
#healthcare
#patient

AI has high requisites to be accepted

Existing systems are not compatible with our AI-initiatives

Lacking in knowledge. What can AI do for us?

Next-gen healthcare systems?

Template-education, what can you achieve with AI?

Inspired from all of the educations caused by the GDPR regulation

New and successful competitors

Fig. 18: Difficulties regarding strategy and technology.

Conclusion. What are our future plans?

In addition to continuing to explore machine learning for the rest of the project period (2018), we already have thoughts for specific side projects we intend to start if we find the right project partners.

We are thinking of moving in one or more of the following directions:

1. **Register research + machine learning**.
2. **Predicting re-admission.** A direction that Sahl-grenska University Hospital e-Psychiatry unit has already mentioned.
3. **Transfer learning within computer vision**, for example diagnostic imaging together with radiologists.
4. **Computer vision:** Train our own VOC (Visual Object Classes) on a medical aspect. Do we have unique data in BFR? It may take several months or years if you do not have the right hardware or manage to parallelise on a larger scale.
5. **Deep/transfer learning** to study faces in the search for diagnoses, self-care tips and self-triage in daily

life. It is important to find a good data source and enough computing power.

6. **NLG** (Natural Language Generation)[58]. To describe measurements and other non-text with plain text or have it read out. For example, it is a challenge for those with poor eyesight to read tabular data, or for people who have comprehension problems a descriptive summary/conclusion may help instead of being forced to burden their working memory.

7. **NLP to study patient accounts/medical history**, perhaps in directions such as prioritising in healthcare. Here, we are already in contact with a doctor at Sahlgrenska University Hospital who we can work with.

8. **NLP to automate the creation of personal compilations of patient information** from sources such as 1177.se etc. This advice could be requested through a conversation-based interface and you can choose to have the compiled material read out or recorded as an audio book.

9. **NLP: Investigate whether word vectors work** as technology in a healthcare context.

10. **Miscellaneous**: A compilation of medically useful, neural networks as a service, downloadable ML models that have already been trained as well as feature detectors. Are there any credible marketplaces for this?

In addition to finding project partners, we will search for funding by replying to suitable calls for proposals.

Appendix

Not everything is suitable for inclusion in the report. Therefore, here are some more details for us geeks.

Glossary

Keep in mind that some of the words in the glossary have a different meaning in other contexts.

Adaptive Boosting (AdaBoost) – a meta-algorithm for finding the most likely response from a multitude of options.

Algorithm – a type of recipe or method which a machine can repeat again and again.

Amazon Echo (Alexa) – a smart speaker you can talk to, ask questions and ask to record things.

Apple Siri, SiriKit – Apple platform for talking to its gadgets. Uses Apple servers in the background to understand what is being said and the user's intention.

Application Programming Interface (API) – a way of talking to anything technical. The interface of a service you hire on the internet for image recognition, for example.

Artificial Intelligence (AI) – the idea that something artificially created can show intelligent behaviour.

Batch learning – the concept that neural networks learn new things intermittently with packages of new material to train on. Compare to online learning.

Bias – to have a slant in some direction. Used about data sources and the conclusions reached by machine learning. One example is the Google Image search that was found to be prejudiced in searches for black teenagers by showing them in a criminal context, while white teenagers smiled and did sports.

Bild- och FunktionsRegistret, BFR (Image and function register) – globally unique medical register including X-ray images etc. from X-ray examinations. Operated by Region Västra Götaland (VGR).

Black box – describes the part of a neural network whose workings are difficult or impossible to explain.

Business Intelligence (BI) – collective term for work to better understand your own business. This is often done by finding data and making decisions based on the data instead of on gut feeling/experience.

Cascading classifiers – see Adaptive Boosting.

Code system – a way to create structure by having codes that correspond for example to a diagnosis, a lab test etc. See International Classification of Primary Care for example.

Cognitive services, cognitive computing – the mental process of acquiring knowledge and understanding through thinking, experiencing and using the senses. Some suppliers call their services cognitive rather than intelligent or provide them with an AI label.

Computer vision (CV) – giving machines the ability to see and interpret visual things such as images, videos etc.

Convolutional Neural Network (CNN) – most often used in computer vision and is a neural network which refines its impression of an image, for example, for each layer. It is therefore a description of how cooperation works between the network's hidden layers.

Decision tree – a structured way of making decisions. Think of a tree trunk as the entrance and, depending on the question, some of the thicker branches become relevant. The more information, the closer to a leaf you get. The leaf can symbolise a diagnosis.

Decision-making support – getting support to reach a decision based on a foundation of data.

Deep learning (DL) – a special version of machine learning which uses many hidden layer and greater complexity than would otherwise be the case.

Entropy – a quantitative measure of how ordered/structured a quantity of information is.

Feature detection/engineering – the engineering part is about developing a way of recognising something, for example an eye in an image. A "detector" is the result of engineering, i.e. something which recognises something in particular. Feature detection is the activity of trying to identify one or more things.

Generative Adversarial Networks (GAN) – two neural network that learn together. One has the role of challenging the other. The result is a network that is good at distinguishing fake from real material and another network that is good at trying to deceive.

GitHub – developer service you can find at github. com, where source code is available, for example for solutions for anything technical.

Google Home – Google's smart speaker, see also Amazon Echo.

GPU (Graphics Processing Unit) – a graphics processor or graphics card. They are extra suitable for machine learning since they can work with many things at the same time.

Hidden layer – normally a number of layers in a neural network. They are called hidden because they are not visible in the same way as the input and output layer. Hidden layers are what is meant when we talk about an AI "black box".

Hypothesis – an assumption of reality, that we think that something is a certain way. For example that this report believes that machine learning is useful for assessing patients' health.

ICD-10 – see Code system.

Input (layer) – input is what you put into a neural network, for example an image; the input layer is the first layer that receives the image before it "disappears" among hidden layers.

Interface – a way to interact with something. Graphical User Interface (GUI) is the one on your mobile phone screen, when you talk it, it is a voice-based interface. When your mobile vibrates with a notification, it is a tactile/haptic interface.

International Classification of Primary Care (ICPC) – one of many code systems offering codes for classification of different disorders in primary care. Means that diagnoses do not need to be written in free text. If you have taken a blood test, the code is "-34", while fever has the code "A03".

KVÅ, Klassifikation av vårdåtgärder (Classification of care measures) – the codes used in reports to the Swedish National Board of Health and Welfare's health data records. Used to compile statistics on health and medical care measures.

Latent Semantic Analysis (LSA) – used to calculate the relationship between words or terms. A way to allow a machine to detect patterns, such as that doctor and physician are in many cases synonymous. But it is also a way of finding out what does not seem to go together, for example that nails and abdomen in an anatomical context are not especially related.

Linked data, Linked Open Data (LOD) – a way to structure information to a level that is closer to the knowledge a machine can absorb. Compare this to how a human being explores Wikipedia.

Machine intelligence (MI) – a machine showing intelligent traits or helping to provide insights.

Machine learning (ML) – a machine learning something without being explicitly instructed in detail.

Machine-readable – information that a machine is being able to absorb. This may prove to be a problem when cognitive ability is required to fill in gaps or a document format shows the content correctly but it has been stored higgledy-piggledy. See also Linked data.

Medical history – account of someone's medical history, a bit like an interview that is written down by health professionals in a meeting with the patient. A controlled conversation to find appropriate measures.

Model, knowledge model, Machine learning model (ML model) – the knowledge that is packaged in a way that allows other machines to learn. Sometimes, the neural network architecture is also included in the model, such as the number of hidden layers, the order of specialised layers etc.

Named Entity Recognition (NER) – a technique in natural language processing for working with text, for example to identify if a person is mentioned in the text, if a diagnosis is mentioned or the dosage of medicine.

Natural Language Generation (NLG) – using technology to create a text or spoken words based on information. This can be a summary that is written/spoken based on collected data.

Natural Language Processing (NLP) – umbrella concept for processing text and speech to understand the content. See also NER, NLTK, NLU etc.

Natural Language Toolkit (NLTK) – technical framework to help with the work of NLP.

Natural Language Understanding (NLU) – getting a machine to understand natural language and speech.

Neural network, Artificial Neural Network (ANN) – an artificially created network of neurons that imitates a brain. Consists of layers such as the input layer, hidden layers and output layer.

Neuron – may also be referred to as brain cells or nerve cells.

Online learning – unlike batch learning, learning in the neural network is continuous. When someone says their AI is constantly improving, it is online learning they are talking about.

Output (layer) – output is the answer itself and the output layer is the layer of a neural network where the answer comes out.

Overfitting vs underfitting – overfitting is drawing incorrect, far-reaching conclusions based on the data. A technical equivalent of bias. Underfitting is not seeing patterns that are actually there. See also Bias.

Patient account – free narrative of the patient's health conditions, unlike the more structured interview that takes place to establish medical history.

Percent, percentage point – percent stands for a change that is relative to an initial value, while percentage points are a change in percentage. For example, if a political party election result decreases from 25% to 20% of the votes, they have lost 20% of their voters (five twenty-fifths), but at the same time they have lost 5 percentage points.

Personas – archetypes of the users we approach with a solution. They often have a name and a list of descriptive characteristics to remind us who the end users are.

Quality register – data collected by the healthcare provider to follow up on the quality of healthcare. Examples of quality registers are the cause of death register, others are about cancer treatment etc.

Region of Interest (ROI) – a limited area in an image, the location of something interesting or worthy of focus.

Reinforcement learning (RL) – a variant of machine learning where the machine can try to learn by knowing what is a desirable result (known as a reward function), as well as what it should try to avoid (known as the cost function). Compare this to cycling without falling off your bike.

Second opinion – to get another opinion on something. Such as requesting a new assessment of X-ray images. This may be possible to automate.

Self-reinforcement – ending up in a cycle of positive/negative reinforcement without external/new impressions. For example if machine learning generated data which it then input again as its own view of the world.

Sentiment analysis – a NLP technique for reading feelings from a mass of text or something that is heard.

SnoMED-CT – see Code system.

Supervised learning (SL) – training a neural network where the data used as a knowledge base also contains answers (known as labels). In the context of healthcare, this includes medical history and diagnostic codes from a doctor. See also Unsupervised learning.

Synapse – connects the brain cells/neurons and enables them to communicate as a network.

Target groups – see Personas.

Toy problems – in AI, the term is used for solutions to banal problems or something of which the benefit is difficult to see.

TPU (Tensor Processing Unit) – a specialised processor from Google designed to be good at machine learning. See also GPU.

Transfer learning (TL) – learning lessons from one area and applying them to something else. For example, neural networks trained on flowers, animals etc. proved to have a certain pre-understanding which can be used in human medicine.

Triage record – the control questions, answers and observations needed to perform triage. This can involve taking the pulse, observing breathing or wounds, for example.

Triage, self-triage – triage is assessing and prioritising how urgent something is. In the Emergency Department, triage is needed to usher in the really serious cases, while getting the less urgent cases to sit in the waiting room. If all knowledge of triage could be digitised, people could perform triage themselves or for a relative.

Turing test – a way of testing if a machine can manage to make a person it talks to believe that the machine is also a human being.

Unsupervised learning (UL) – unlike supervised learning, this has a mass of data but no answers. This means looking for structure in data, classifying and grouping it in the search for something useful.

Validation data, training data – to train a machine, the data source is divided into at least two different groups. Training data is used to train the machine and is normally the bulk of the original data source. In order to be able to assess how accurate the machine is at predicting something, some validation data that the machine does not know anything about is saved. When the machine is then supplied with new data, it is possible to measure how well it performs – how good its "fitting" is. See also Overfitting.

Visual Object Classes (VOC) – describe visual objects such as cars, people, meatballs, difficult-to-heal scrapes and birth marks which need to be treated. Where the non-medical is concerned, there is a visual equivalent to the Turing test, a competition to see if a human being or a machine is most accurate at identifying objects.

Vitalis – annual conference in Gothenburg on e-health, healthcare technology etc.

Weak classifiers – factors which do not individually have a large impact on the result, but which together can decide what something is. For example that pointy ears, whiskers and low wither height means it is more likely to be a cat than a dog.

WEIRD – a term in psychology, is an abbreviation of Western, educated, industrialised, rich and democratic. Points out the challenge that those who design algorithms do not always have a lot in common with all their users.

Word vectors – a newer solution for what Latent Semantic Analysis is trying to solve. Word vectors involve clarifying linguistic issues such as which words seem to belong together, for example phrases, proverbs

etc. Word2Vec is one such solution which has been
released for free by Google.

Further material

Articles

450,000 Women Missed Breast Cancer Screenings Due to "Algorithm Failure"
spectrum.ieee.org/riskfactor/computing/
it/450000-woman-missed-breast-cancer-screening-exams-in-uk-
due-to-algorithm-failure

AI winter is well on its way
Criticism of those who oversell the progress made in AI
blog.piekniewski.info/2018/05/28/ai-winter-is-well-on-its-way/

Using electronic patient records to discover disease correlations and stratify patient cohorts
About seeing the correlation and interaction between various
medications
www.ncbi.nlm.nih.gov/m/pubmed/21901084/

Patient stratification and identification of adverse event correlations in the space of 1190 drug related adverse events
www.ncbi.nlm.nih.gov/m/pubmed/25249979/

Data is not the new oil
techcrunch.com/2018/03/27/data-is-not-the-new-oil/

IBM outlines the 5 attributes of useful AI
venturebeat.com/2018/04/21ibm-outlines-the-5-attributes-of-
useful-ai/

Is AI Riding a One-Trick Pony?
That without back propagation we would not care about AI
today

www.technologyreview.com/s/608911/
is-ai-riding-a-one-trick-pony/

The replication crisis
Difficult to reproduce/replicate findings in AI/ML
www.sciencemag.org/news/2018/02/
missing-data-hinder-replication-artificial-intelligence-studies

How to jump start your deep learning skills using Apache MXNet
About deep learning and NLP
www.oreilly.com/ideas/how-to-jump-start-your-deep-learning-skills-using-apache-mxnet

Algorithm can predict if you'll live after a heart transplant with scary accuracy
The algorithm is 14% better than state-of-the-art knowledge
www.newsweek.com/heart-transplant-organ-donation-tree-predictors-934204

Why is machine learning 'hard'?
As a result of having two additional layers to troubleshoot compared to other software development
ai.stanford.edu/~zayd/why-is-machine-learning-hard.html

300 000 gånger mer data i AI-modeller (300,000 times more data in AI models)
Specialised processors (GPUs and TPUs) make it possible to perform calculations on more data, the amount has doubled every 3.5 months since 2012 according to the article
techworld.idg.se/2.2524/1.702799/mer-data-ai

Word2vec
A solution for finding out what relationships are contained in text, such as which words are included in phrases, proverbs or other types of pattern, e.g. that first and last names often follow each other in a text
en.wikipedia.org/wiki/Word2vec

Grandmother cell
We have specialised brain cells, e.g. for quickly recognising our grandmother
en.wikipedia.org/wiki/Grandmother_cell

The Guardian view on AI in the NHS: a good servant, when it's not a bad master
Perspective on data collection
www.theguardian.com/commentisfree/2018/may/21/
the-guardian-view-on-ai-in-the-nhs-a-good-servant-when-its-
not-a-bad-master

A Guide to Successful AI Implementations, and Why So Many Fail
How to set up an AI project and learn from those who have failed.
narrativescience.com/Resources/Resource-
Library/Article-Detail-Page/a-guide-to-
successful-ai-implementations-and-why-so-
many-fail

Gender shades - How well do IBM, Microsoft, and Facebook AI services guess the gender of a face?
gendershades.org

FairML: Auditing Black-Box Predictive Models
Being able to inspect ML models
blog.fastforwardlabs.com/2017/03/09/fairml-
auditing-black-box-predictive-models.html

Books

Machine Learning Yearning
by Andrew Ng
www.mlyearning.org

Data Science for Business: What you need to know about data mining and data-analytic thinking
by Foster Provost and Tom Fawcett

www.goodreads.com/book/show/17912916-data-science-for-business

Our Final Invention: Artificial Intelligence and the End of the Human Era
by James Barrat
www.goodreads.com/book/show/17286699-our-final-invention

Technical solutions

Apache MXNet
Scalable environment for learning so data does not leave the organisation
mxnet.apache.org

fitchain.io

FairML
Framework in Python for finding the bias in machine learning
github.com/adebayoj/fairml

IBM Bluemix
Partly supports even Swedish in NLP – keywords, metadata and entities
console.bluemix.net/docs/services/natural-language-understanding/language-support.html#swedish

Video clip, demo & podcast section

2D Visualization of a Convolutional Neural Network
Test drive solution for recognising numbers
scs.ryerson.ca/~aharley/vis/conv/flat.html

Hello World - Machine Learning Recipes #1
Code your own machine learning in approximately five minutes
www.youtube.com/watch?v=cKxRvEZd3Mw

Neural Network 3D Simulation
Visualisation of image interpretation in multiple layers of a
neural network
www.youtube.com/watch?v=3JQ3hYko51Y

Reproducibility and the Philosophy of Data
with Clare Gollnick in TWiML Talk
overcast.fm/+GooiXS37s

The no free lunch theorems
by Data Skeptic
overcast.fm/+BzxPVrg7k

Data sources etc.

SND - Svensk Nationell Datatjänst (Swedish National Data Service)
For finding your way in the available research data
snd.gu.se/sv

Öppet API för Försäkringsmedicinskt Beslutsstöd
From the Swedish National Board of Health and Welfare
api.socialstyrelsen.se/fmb/dokumentation/psi/swagger-ui.html

Watson Data Kits
Pre-cleaned and structured data from IBM in a number of
different industries
console.bluemix.net/catalog/services/watson-data-kits

PASCAL VOC (Visual Object Classes)
Examples of visual pattern recognition
host.robots.ox.ac.uk/pascal/VOC/index.html

High Performance Computing (HPC) in the Cloud
For hiring computing power as a service
aws.amazon.com/hpc/

The project's interim reports in the developer blog

Interim reports in chronological order from the developer blog. All in Swedish.

AI-projekt under uppstart –inventering av produktifierade tjänster
We started by testing what the suppliers offer
vgrblogg.se/utveckling/2018/02/13/ai-projekt/

AI: Naturligt språkprocessering förstärkt av svartkonst
About NLP
vgrblogg.se/utveckling/2018/03/05/naturligt-sprakprocessering-nlp-forstarkt-av-svartkonst/

AI-projektet: Prototyp på app för smartklocka
Example of demonstrator of end solution
vgrblogg.se/utveckling/2018/03/21/prototyp-pa-app-for-smartklocka/

AI-projektets informationsarkitektur
About the data source we inspected with medical history and ICPC codes with diagnoses
vgrblogg.se/utveckling/2018/04/22/ai-projektets-informationsarkitektur/

Att konversera med maskiner
How much of what you say do these machines understand?
vgrblogg.se/utveckling/2018/04/25/att-konversera-med-maskiner/

GAN tränar en maskins förståelse genom att duellera sig själv
Letting a machine learn by itself
vgrblogg.se/utveckling/2018/05/01/gan-tranar-en-maskins-forstaelse-genom-att-duellera-sig-sjalv/

Deep learning till beslutsstöd för att klassificera tecken på stroke?
Classification of a facial expression
vgrblogg.se/utveckling/2018/05/03/deep-learning/

AI och computer vision för att bli mer redo för GDPR
About the classification of objects
vgrblogg.se/utveckling/2018/05/07/
ai-och-computer-vision-for-gdpr/

Idéworkshop om möjligheter, svårigheter och lösningar
Workshop where different people who work in the healthcare sector contributed
vgrblogg.se/utveckling/2018/05/28/ai-ideworkshop/

Thanks to

The proofreaders and those who contributed tips on content. Particular thanks to:

- Agneta Grangård
- Kerstin Hinz
- Almira Thunström
- Stuart Filshie
- Kristian Norling
- Martin Adiels

Image sources

- www.commitstrip.com/en/2017/06/07/ai-inside/
- aiweirdness.com/post/171451900302/do-neural-nets-dream-of-electric-sheep
- xkcd.com/1838/
- twitter.com/scanlime
- twitter.com/AgnesWold
- Cover image by starline / Freepik.

Source references

1   sv.wikipedia.org/wiki/Artificiell_intelligens

2   sv.wikipedia.org/wiki/Maskininlärning

3   sv.wikipedia.org/wiki/Sjukdomshistoria

4   en.wikipedia.org/wiki/Natural-language_
processing

5   en.wikipedia.org/wiki/Named-entity_
recognition

6   en.wikipedia.org/wiki/International_
Classification_of_Primary_Care

7   sv.wikipedia.org/wiki/SNOMED_CT

8   sv.wikipedia.org/wiki/ICD-10

9   www.socialstyrelsen.se/klassificeringochkoder/
atgardskoderkva

10   en.wikipedia.org/wiki/Natural_Language_
Toolkit

11   sv.wikipedia.org/wiki/Deep_learning

12   sv.wikipedia.org/wiki/Länkade_data

13   api.socialstyrelsen.se/fmb/dokumentation/psi/swagger-ui.html

14   www.dyslexiforeningen.se/vad-ar-dyslexi/

15   en.wikipedia.org/wiki/Feature_detection_(computer_vision)

16   en.wikipedia.org/wiki/Convolutional_neural_network

17   sv.wikipedia.org/wiki/Triage

18   en.wikipedia.org/wiki/Decision_tree

19   en.wikipedia.org/wiki/Weak_AI

20   computersweden.idg.se/2.2683/1.698148/it-kopare-ai

21   en.wikipedia.org/wiki/Turing_test

22   en.wikipedia.org/wiki/Business_intelligence

23   en.oxforddictionaries.com/definition/intelligence

24   en.wikipedia.org/wiki/How_to_Lie_with_Statistics

25   en.wikipedia.org/wiki/Artificial_neural_network

26   www.scientificamerican.com/article/one-face-one-neuron/

27   en.wikipedia.org/wiki/Supervised_learning

28   en.wikipedia.org/wiki/Labeled_data

29   en.wikipedia.org/wiki/Unsupervised_learning

30   en.wikipedia.org/wiki/Reinforcement_learning

31   en.wikipedia.org/wiki/Transfer_learning

32   en.wikipedia.org/wiki/Kullback–Leibler_divergence

33   en.wikipedia.org/wiki/Diversity_index

34   en.wikipedia.org/wiki/Online_machine_learning

35   en.wikipedia.org/wiki/Sentiment_analysis

36	www.chalmers.se/sv/institutioner/cse/
kalendarium/Sidor/Thesis-Defence-Olof-Mogren.aspx

37	en.wikipedia.org/wiki/AdaBoost

38	en.wikipedia.org/wiki/Cascading_classifiers

39	en.wikipedia.org/wiki/Latent_semantic_analysis

40	en.wikipedia.org/wiki/Latent_semantic_
analysis#Applications

41	developer.infermedica.com/docs/nlp

42	www.gu.se/omuniversitetet/personal/?
languageId=100000&disableRedirect=true&
returnUrl=http%3A%2F%2Fwww.gu.se%2Fenglish
%2Fabout_the_university%2Fstaff%2F%3
FlanguageId%3D100001%26userId%3Dxrawar&
userId=xrawar

43	aiweirdness.com/post/171451900302/do-neural-
nets-dream-of-electric-sheep

44	twitter.com/picdescbot

45	en.wikipedia.org/wiki/Boosting_(machine_
learning)#Boosting_algorithms

46	en.wikipedia.org/wiki/Feature_engineering

47	scapis.se/

48	techworld.idg.se/2.2524/1.703287/nvidia-hgx2

49	well.blogs.nytimes.com/2016/03/14/hey-siri-can-
i-rely-on-you-in-a-crisis-not-always-a-study-finds/

50	vgrblogg.se/utveckling/2016/09/09/
jamlik-vard-i-algoritmernas-varld/

51	rationalwiki.org/wiki/WEIRD

52	spectrum.ieee.org/riskfactor/computing/
it/450000-woman-missed-breast-cancer-screening-
exams-in-uk-due-to-algorithm-failure

53	theflatearthsociety.org/

54	vgrblogg.se/utveckling/2018/03/21/prototyp-pa-

app-for-smartklocka/

55   github.com/Vastra-Gotalandsregionen/health-guide-for-apple-watch

56   github.com/marcusosterberg/triage-at-home/blob/master/Triage-at-home.ipynb

57   github.com/marcusosterberg/triage-at-home

58   en.wikipedia.org/wiki/Natural_language_generation

VÄSTRA
GÖTALANDSREGIONEN

SWElife