AMPS-QT

The AMPS Quarterly - Issue n.42 - 2Q2019 Editors: Fabio Badilini PhD, FACC and Martino Vaglio MS

AMPS-QT is a quarterly journal dedicated to all the people and organizations involved in the world of cardiac safety. Published by AMPS LLC, it covers all aspects of methodology and software technology related to clinical trials and Thorough QT studies.

Editorial

Recently one of our customers asked if AMPS plans to develop Artificial Intelligence powered products. We can tell the truth now, AMPS has been using AI (in today terminology) in some of our tools for several years, but never really promoted the fact in our communications as AI is very much the buzzword. Too often the term gets distorted to include any kind of algorithm or computer program. We thank Silicon Valley (and, cough, Mark Zuckerberg, cough) for constantly inflating the capabilities of AI for its own gratification.

So, we thought it is the right time to provide our readers some basic information about AI and explain AMPS' position in respect to AI technology and how it impacts our industry from our point of view. In this issue of AMPS-QT you will find, a short introduction to AI, followed by AMPS' assessment of the current state of affairs, and finally a Q&A session about the effectiveness of AI in ECG Signal Processing with Roberto Sassi and Matteo Bodini, of the University of Milan. Please enjoy this presentation and feel free to contact AMPS with your ideas, reactions, and opinions on and about AI.

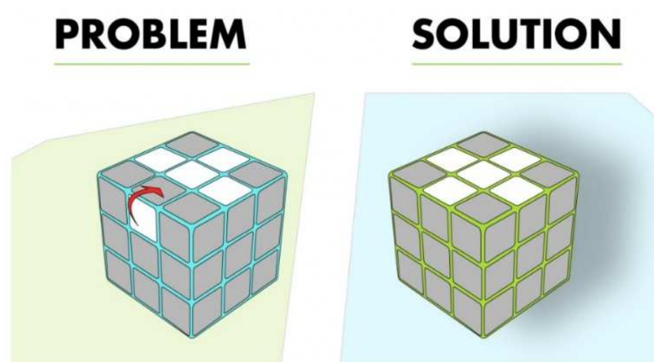
AMPS Views on:

AMPS and AI

By Fabio Badilini, PhD, FACC, AMPS LLC.

What is AI, exactly? The question may seem simple, but the answer is complicated. In the broadest sense, AI refers to machines that can learn, reason, and act for themselves. AI machines are supposed to be able to solve problems and make their own decisions when faced with new situations, in a manner similar to that of humans and animals.

So, how do you develop AI? To see an example, let's take a simple step by step description of problem solving. The figure below depicts a fairly simple problem: how to move a square on a Rubik's cube from a position A to a position B.



A human proficient at puzzles, using natural intelligence, determines how to solve the problem, and describes the solution with this set of instructions:

1. **F**ront inverse (counterclockwise 90 degrees)
2. **U**p (clockwise 90 degrees)
3. **L**eft inverse (counterclockwise 90 degrees)
4. **U**p inverse (counterclockwise 90 degrees)

Or **F' U L' U'** for people familiar with solutions to a Rubik's Cube puzzle.

The instructions above are an example of the basis for what is commonly called an ***Algorithm***.

An Algorithm is therefore nothing but a process, or set of rules, to be followed in a correct sequence for calculations or other problem-solving operations. Of course, to be executed by a computer, these rules need to be written in a programming language, so the algorithm gets imbedded into a program which then, given a set of data as input, produces the solution.

Depending on the complexity of the task at hand therefore Algorithms can be very simple, such as, solve the Rubik's Cube problem, or quite complicated, for example, play a game against the best human chess champion, like Deep Blue, the algorithm designed and developed by IBM that defeated Grand Master Garry Kasparov in 1996. Dependent on the complexity of the process to be performed by the Algorithm it has become, in modern terminology, to name some of these Algorithms as AI.

From Algorithm to AI

AI is therefore software that can 'reason' on input, follow a process, and explain on output. There seem to be consensus in defining three major AI categories: ANI, AGI, and ASI.

ANI stands for **Artificial Narrow Intelligence**. It's often called Weak AI. ANI/AI is referred to as weak because it specializes in only one category (example: directing your car from A to B via the fastest route). AGI, or **Artificial General Intelligence**, is also called Strong AI or Human-Level AI. As the name infers, AGI is supposed to have the same capabilities as a human. ASI, or **Artificial Superintelligence**, is the type of intelligence that is smarter than humans. Today's world is running on ANI, and AGI may be created in the near future, while ASI is still in the distant future (although people believe that when AGI is operative it will develop immediately into ASI).

For now, let's focus on ANI. As we said, weak artificial intelligence is designed to be focused on a narrow task and to seem very intelligent at it. Paraphrasing what we wrote years ago in AMPS-QT however: *"One should not rely solely on a computer-generated solution It's only as accurate as the human who programmed the machine!"*. This is still true today. In turn, the value of any specific ANI is very much dependent on how 'smart' were the humans who developed the algorithm.

By this definition most of the AMPS products are already based on ANI, in particular those using the BRAVO and ABILE algorithms, such as, FAT-QT and CER-S. These algorithms analyze the data and are able to determine for example, the length of a QT interval, or recognize the pattern of an Atrial Fibrillation, and other factors. It took a team-effort years to define and refine these algorithms. As a result, their superior effectiveness has been demonstrated against similar commercial products from various ECG manufacturers and documented in several peer-review articles. Are the AMPS ANI-based products perfection? Of course not, no ANI is 100% perfect. But that is not from a lack of effort, each year AMPS invests a majority of the company's profit into research and development. With our commitment, we are optimistic AMPS' ANI can reach perfection in the future!

In today's world of AI, most of the ANI advancements and applications you hear publicized refer to a category of algorithms known as 'machine learning'. These algorithms use statistics to find patterns in massive amounts of data. They then use those patterns to make predictions on things like what shows you might like on Netflix, or which book or music you're likely to buy when you connect to Amazon. Machine learning, and its subset 'Deep Learning' (basically machine learning on steroids), are incredibly powerful. They form the basis of many recent breakthroughs, including facial recognition and hyper-realistic photo and voice synthesis.

AI and ECG Signals

The question now is: What is the possibility to improve ANI using Deep learning techniques for a better interpretation of ECG and Holter traces, and for the benefit of the cardiac safety industry at large?

The literature on the topic is already extremely rich and dates to more than 2 decades ago [1]. However, apart from extremely specific applications, a breakthrough is yet to be made. In some (alas very frequent) cases, the selection of the right neural network model was not adequately contemplated. Even worse, although not surprisingly, the choice of the learning (training) and of the validation (test) sets were driven by the 'easiest' choice, namely, the use of standard solutions from public domain datasets (namely the MIT PhysioNet records). These datasets were established many years ago for totally different goals, and which (in most scenarios) do not represent the ideal choice for the specific problem. It is not infrequent to see extremely complex networks (designed for another purpose like facial or speech recognition) being trained with totally unsuited learning sets. To the point that one could ironically name this "small data" AI.

One noticeable exception is the separation between sinus rhythm and atrial fibrillation rhythms from short recordings, for which, and also thanks to the push of the wearable industry (AliveCor, and more recently Apple) the added value of an AI approach based on big data has been significantly demonstrated.

A few recent publications have shown a change in the trends of usage. In the first of its kind, and perhaps a most remarkable paper published in early 2019 on Nature, Dr. Hannun and colleagues (from a joint collaboration between Stanford and UCSF universities) were able to discriminate between 12 different heart rhythm patterns using a single lead patch device [2]. In this study, a complex convolutional neural network was trained using a set of 91,232 ECG records from 53,549 patients (mean wear time 10.6 days) and the results over a much smaller test set of 328 records were remarkably encouraging. In fact, superior to that of a team of board-certified practicing cardiologists.

The path for AI-applied ECG waveform analysis is still at its early phase. We believe a mixed model, where the power of 'deep learning AI' could be combined with the 'intelligence' gathered by traditional signal processing will provide the ultimate solution, for both arrhythmia and ECG segmentation measurement algorithms.

Q&A on Artificial Intelligence in ECG Signal Processing

Matteo Bodini, MS & Roberto Sassi, PhD; University of Milan, Milan, Italy

Q1. How would you summarize the state of the art of Artificial Intelligence applied to ECG signal processing?

Artificial Intelligence (AI) has been widely applied to ECG signals. Mainly, AI has been exploited for the tasks of disease classification, emotion recognition, and biometric identification. With respect to the first, arrhythmias classification was the most tackled goal.

Two frameworks were mainly implemented. The first is the classical AI pipeline: handcrafted features extraction, followed by classification using well-known algorithms (K-Nearest Neighbors, Support Vector Machines, Random Forests, and others). The second is represented by the newest Deep Learning (DL) models, where features are learned from data. Handcrafted features carry several problems: they have to be manually designed, they are sensitive to noise and artifacts of ECG signals, and further, they imply high-dimensionality feature spaces, as usually many features are required for classifying a given condition. Dimensionality reduction is easier to achieve through DL, but the results lack any kind of 'explainability', that is, DL models often result in better classification performances, but researchers are almost always unable to understand their classification decisions.

Regarding the datasets, an effort has been put in working on large collections of ECG signals. However, the datasets which were most used (such as, the ones publicly available on PhysioNet) are often of small size. Also, single or two lead signals were mostly considered. They are adequate to feed DL models only for the task of beats classification, where a large number of samples can be easily obtained from each signal. However, building models for detecting arrhythmias that can be diagnosed through the simultaneous analysis of multiple leads is difficult with the datasets available today.

In conclusion, the advancements of AI in ECG signal processing are remarkable. However, in our opinion, two issues must be addressed in the near future:

1) 'explainability' of AI techniques, that is, explain how AI methods (in particular DL) take decisions;

2) collecting bigger datasets to try to extend AI classifiers at the level of current ECG analysis programs (as a reference, in image classification problems millions of images are considered, so it is likely that many millions of ECG records might be necessary).

Q2. Do you think that the choices made with respect to the selection of the neural network model and the training/learning sets used, are generally adequately assessed on the basis of the specific working hypothesis, or tailored to the clinical problem that one is trying to solve?

Within the world of AI, the choice of the specific model (either among classical techniques, such as Support Vector Machine, K-Nearest Neighbors, Random Forests, or deep learning networks) and the choice of the train/test sets are often based on experimental verification: in general, no theoretical rule is available, given a specific problem.

Let's now focus on AI applied to ECG signals and in particular to the task of beat classification. In our opinion, the models are almost always selected in a fair way. That is, the best model is chosen according to (quantitative) performance metrics (typically accuracy, when classes are balanced, or other measures as precision and recall when they are not). We do not feel that the same can be always said for the choice of the train and test sets. A wrong selection can cause overfitting, that is, the model works well on training data but gets worse performances on unseen test data. Train and test sets are often chosen with a ratio of (approximatively) 70:30. While this is not a problem with very large datasets, with smaller ones the distribution of the data, contained in each, is critical in evaluating the performances. A random split may return a situation from which the model can easily get good performances, or, the opposite, only bad ones.

Many statistical techniques are available for partially overcoming these issues. The most popular one is k fold cross-validation: the entire dataset is divided into k folds, then the algorithm is applied k times jointly using k-1 folds as train set and the remaining one as a test set. Averaging the performances of the k executions, we obtain a quite fair estimate of the performance of the algorithm, when further applied on other data. The splitting value is often set to k=10, a value determined empirically as effective.

Summarizing, in the world of AI applied to ECG signal processing, models are carefully selected, but the sizes of train and test sets, as well as the distribution of the data in each one, are not widely addressed topics. K fold cross-validation is almost never used (especially with DL models) due to computational requirements (DL models are already computationally demanding when ran once, so repeating the process k times is often practically unfeasible). Even in this case, much larger datasets will mitigate these issues, providing the possibility of selecting ample train and test sets, where the natural distribution of the data (also of less represented conditions) will be respected. However, it is still unclear if this will be a viable solution also for those rare ECG conditions, which are nevertheless correctly diagnosticated by current ECG analysis programs.

Q3. Images or single-dimension vector? For an old-fashioned engineer the idea to use a bi-dimensional representation may seem shocking. One should at least define to better characterize the image format and resolution. What is your opinion?

It is a possible choice, in particular with deep learning models, to stacker the different leads (or beats) in a matrix (example: each row a lead) and then use the matrix in a standard AI pipeline, as it were an image (AI models for object recognition in images are more advanced and historically received a larger attention). While this might seem shocking, it is, at the end of the day, not that different from building a matrix from the ECG leads to subsequently derive its principal components through singular vector decomposition. This technique has been used extensively in biomedical signal processing, for example in the context of atrial fibrillation (to extract the fibrillatory wave). So, image or single-dimension vector shouldn't be considered differently in this context.

References

[1] Shishvan OR, Zois D-S, Soyata T. Machine Intelligence in Healthcare and Medical Cyber Physical Systems: A Survey. IEEE Access 2018, Vol.6:46419-46494.

[2] Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med. 2019 Jan; 25(1):65-69.

Products News

Latest Releases

In 2019Q2, we released an updated version of CER-S (v 3.2.0), including the following revised platforms:

- Continuous ECG beat detection and classification, including the fully renewed ABILE algorithm, with new long analysis capability, up to 30 days.
- ECG beat editor
- Arrhythmia detection and Arrhythmia editor
- ECG Beat Measure, for measuring averaged time-templates ECG complexes, including ST-displacement assessment.
- Report generation.

Looking forward

In 2019Q3, we are going to release the version 4.0.0 of CER-S, with all platforms revised and with the addition of the Events Review. This version will have the CE certification for medical device.

In late Q3 or early Q4, CER-S will also be certified by the FDA, via 510K.

AMPS Notebook

Fabio attended the **44th ISCE Annual Conference** held in April in Atlantic Beach, Florida.

In May, Fabio also attended the **40th** Annual **Heart Rhythm** Scientific Sessions held in San Francisco, California.