Learning signals and systems and its applications to electrocardiography

Antônio Horta Ribeiro

Presented to Aalto community as part of the application for the assistant professor position in Al-Health

Aalto University, Finland (Online) June 2022

Example: modeling deepwater oil well



Figure: Offshore extraction

- Example: modeling deepwater oil well
 - $y_t \rightarrow$ Pressure at the bottom;



Figure: Offshore extraction

- **Example:** modeling deepwater oil well
 - $y_t \rightarrow$ Pressure at the bottom;
 - $u_t \rightarrow$ Pressure & temperature at platform;



Figure: Offshore extraction

- **Example:** modeling deepwater oil well
 - $y_t \rightarrow$ Pressure at the bottom;
 - $u_t \rightarrow$ Pressure & temperature at platform;
 - $t \rightarrow time;$



Figure: Offshore extraction

- Example: modeling deepwater oil well
 - $y_t \rightarrow$ Pressure at the bottom;
 - $u_t \rightarrow$ Pressure & temperature at platform;
 - $t \rightarrow$ time;
- Auto-regressive model:





Figure: Offshore extraction

- Example: modeling deepwater oil well
 - $y_t \rightarrow$ Pressure at the bottom;
 - $u_t \rightarrow$ Pressure & temperature at platform;
 - $t \rightarrow$ time;
- Auto-regressive model:



Data-driven methods:

$$\min \sum_t \mathsf{dist}(\hat{y}_t, y_t)$$



Figure: Offshore extraction

- Example: modeling deepwater oil well
 - $y_t \rightarrow$ Pressure at the bottom;
 - $u_t \rightarrow$ Pressure & temperature at platform;
 - $t \rightarrow$ time;
- Auto-regressive model:



Data-driven methods:

$$\min \sum_t \mathsf{dist}(\hat{y}_t, y_t)$$

Long-term predictions:





Figure: Offshore extraction

Learning recurrent models

Can we directly optimize the long-term predictions to obtain better models for the desired use-case?

Learning recurrent models

Can we directly optimize the long-term predictions to obtain better models for the desired use-case?

Noise analysis

A. H. Ribeiro and L. A. Aguirre, "Parallel Training Considered Harmful" Neurocomputing (2018)

Optimization analysis

A. H. Ribeiro, K. Tiels, J. Umenberger, T. B. Schön, and L. A. Aguirre, "On the smoothness of nonlinear system identification," Automatica (2020)

Long-term memory

A. H. Ribeiro, K. Tiels, L. A. Aguirre, and T. B. Schön, "Beyond exploding and vanishing gradients" AISTATS (2020)

Learning recurrent models

Can we directly optimize the long-term predictions to obtain better models for the desired use-case?

Noise analysis

A. H. Ribeiro and L. A. Aguirre, "Parallel Training Considered Harmful" Neurocomputing (2018)

Optimization analysis

A. H. Ribeiro, K. Tiels, J. Umenberger, T. B. Schön, and L. A. Aguirre, "On the smoothness of nonlinear system identification," Automatica (2020)

Long-term memory

A. H. Ribeiro, K. Tiels, L. A. Aguirre, and T. B. Schön, "Beyond exploding and vanishing gradients" AISTATS (2020)

Contribution: Provide throughout analysis of the tradeoffs between extra-flexibility and the optimization challenges these model provide.

Outline

Automatic analysis of the electrocardiogram

Robustness and overparametrization

Research vision

The electrocardiogram (ECG) exam

Cardiovascular diseases:

- leading cause of death globally.
- ▶ \approx 18 million deaths in 2019 (32%).
- ► Myocardial infarction ≈9 million
- 3/4 of them in low/middle-income countries.

The electrocardiogram (ECG) exam

Cardiovascular diseases:

- leading cause of death globally.
- ▶ \approx 18 million deaths in 2019 (32%).
- Myocardial infarction \approx 9 million
- 3/4 of them in low/middle-income countries.

The ECG is the major diagnostic tool.

- Low-cost, safe and non-invasive
- Can detect: arrhythmias, myocardial infarction, cardiomyopathy...





Computational electrocardiography



Figure: Automated ECG interpretation Glasgow Royal Infirmary (1971).

Macfarlane, P.W.; Kennedy, J. "Automated ECG Interpretation-A Brief History from High Expectations to Deepest Networks." Hearts 2021.

Can we improve on existing automatic tools?

A. H. Ribeiro, M.H. Ribeiro, Paixão, G.M.M., D. M. Oliveira, P. R. Gomes, J. A. Canazart, M. P. S. Ferreira, C. R. Andersson, P. W. Macfarlane, W. Meira Jr., T. B. Schön, A. L. P. Ribeiro "Automatic diagnosis of the 12-lead ECG using a deep neural network," Nature Communications, 2020

The Telehealth Center of Minas Gerais



- The Telehealth Center of Minas Gerais
- More than 4000 ECGs per day



- The Telehealth Center of Minas Gerais
- More than 4000 ECGs per day
- Automatic ECG analysis software used as auxiliary tool



- The Telehealth Center of Minas Gerais
- More than 4000 ECGs per day
- Automatic ECG analysis software used as auxiliary tool
- Better accuracy could enable additional uses



Deep neural networks in computer vision



Imagenet

Figure: Models accuracy on ImageNet benchmark. In 2012, LeNet-5 wins ILSVRC.

Deep learning for automatic diagnosis

- ► CODE dataset: historical data 2010 to 2017.
 - ▶ 2.3 M ECGs from n = 1.6 M patients
 - 6 abnormalites

Deep learning for automatic diagnosis

- ▶ CODE dataset: historical data 2010 to 2017.
 - > 2.3 M ECGs from n = 1.6M patients
 - 6 abnormalites
- CODE-test: annotate by 3 cardiologists (n=827).



A. H. Ribeiro , M.H. Ribeiro, Paixão, G.M.M., et al. "Automatic diagnosis of the 12-lead ECG using a deep neural network," Nature Communications, 2020

Deep learning for automatic diagnosis

- ▶ CODE dataset: historical data 2010 to 2017.
 - > 2.3 M ECGs from n = 1.6 M patients
 - 6 abnormalites
- CODE-test: annotate by 3 cardiologists (n=827).
- ► Develop and evaluate deep neural network ← Conv BN ReLU ← ResBik ResBik ResBik ResBik → Dense σ ← Pooling Conv ← Conv BN ReLU Dropout ← Conv BN ReLU Dropout ← BN ReLU Dropout ← BN ReLU Dropout



A. H. Ribeiro , M.H. Ribeiro, Paixão, G.M.M., et al. "Automatic diagnosis of the 12-lead ECG using a deep neural network," Nature Communications, 2020

Results

| | F1 Score | | | |
|-------|----------|---------|--------|-------|
| | DNN | cardio. | emerg. | stud. |
| 1dAVb | 0.897 | 0.776 | 0.719 | 0.732 |
| RBBB | 0.944 | 0.917 | 0.852 | 0.928 |
| LBBB | 1.000 | 0.947 | 0.912 | 0.915 |
| SB | 0.882 | 0.882 | 0.848 | 0.750 |
| AF | 0.870 | 0.769 | 0.696 | 0.706 |
| ST | 0.960 | 0.882 | 0.946 | 0.873 |

Figure: Performance of the deep neural network (DNN) agains experts.

cardio. \rightarrow 4th year cardiology residents

emerg. \rightarrow 3rd year emergency residents

stud. \rightarrow 5th year Medical students

Results



Bootstrapped F1 score values

Figure: Performance of the deep neural network (DNN) agains experts.

cardio. \rightarrow 4th year cardiology residents

emerg. \rightarrow 3rd year emergency residents

stud. \rightarrow 5th year Medical students

Proof of concept study.

- Proof of concept study.
- ► Goal: better telehealth service

- Proof of concept study.
- ► Goal: better telehealth service
- CODEv2 dataset.
 - 2018 2020
 - n = 1M patients
 - Annotated for 60 classes

- Proof of concept study.
- ► Goal: better telehealth service
- CODEv2 dataset.
 - 2018 2020
 - n = 1M patients
 - Annotated for 60 classes
- Different aspects improved in latter studies
 - Interpretability:

- W. Meira Jr, A. L. P. Ribeiro, D. M. Oliveira, and A. H. Ribeiro, "Contextualized Interpretable Machine Learning for Medical Diagnosis," Communications of the ACM, 2020

- D. M. Oliveira, A. H. Ribeiro, J. A. O. Pedrosa, G. M. M. Paixao, A. L. P. Ribeiro, and W. Meira Jr, "Explaining end-to-end ECG automated diagnosis using contextual features," ECML-PKDD, 2020.

Use of images instead of signals:

- V. Sangha, B. J. Mortazavi, A. D. Haimovich, A. H. Ribeiro, C. A. Brandt, D. L. Jacoby, W. L. Schulz, H. M. Krumholz, A. L. P. Ribeiro, Rohan Khera "Automated multilabel diagnosis on electrocardiographic images and signals," Nature Communications, 2022.

Use of ensembles:

A. H. Ribeiro, D. Gedon, D. M. Teixeira, M. H. Ribeiro, A. L P. Ribeiro, T. B Schon, W. Meira Jr "Automatic 12-lead ECG classification using a convolutional network ensemble," CinC, 2020.

Deep learning in electrocardiography

Q1: Can we improve on existing automatic tools?

Q2: Can we use data-driven methods to detect new conditions?

Can we detect non-STEMI cases from the ECG?

Can we detect non-STEMI cases from the ECG?



Diagnosis:

- ▶ ST-elevation MI (STEMI) \rightarrow ECG.
- ▶ non-STEMI \rightarrow blood testing

Left: No MI. Right: STEMI.

Can we detect non-STEMI cases from the ECG?



Diagnosis:

- ▶ ST-elevation MI (STEMI) \rightarrow ECG.
- ▶ non-STEMI \rightarrow blood testing

Dataset:

- Stockholm emergency: 2007 to 2016
- 400 000 ECGs
- ▶ NSTEMI \approx 1%. STEMI \approx 0.5%
- MI annotated using blood testing.

Left: No MI. Right: STEMI.

Can we detect non-STEMI cases from the ECG?

Diagnosis:

- ▶ ST-elevation MI (STEMI) \rightarrow ECG.
- ▶ non-STEMI \rightarrow blood testing

Dataset:

- Stockholm emergency: 2007 to 2016
- 400 000 ECGs
- ▶ NSTEMI \approx 1%. STEMI \approx 0.5%
- MI annotated using blood testing.



Left: No MI. Right: STEMI.



Figure: ROC and precisicion-recall curves AUROC = 0.82 / 0.74 (STEMI/NSTEMI).

Study of bio-markers from ECG to detect mortality

E. M. Lima^{*}, **A. H. Ribeiro**^{*}, G. M. M. Paixão et al., "Deep neural network estimated electrocardiographic-age as a mortality predictor," Nature Communications, vol. 12, 2021 *Equal contribution

Study of bio-markers from ECG to detect mortality

E. M. Lima^{*}, **A. H. Ribeiro^{*}**, G. M. M. Paixão et al., "Deep neural network estimated electrocardiographic-age as a mortality predictor," Nature Communications, vol. 12, 2021 * Equal contribution

Predicting risk of developing Atrial Fibrilation

S. Biton, S. Biton, A. H. Ribeiro, G. Miana, C. Moreira, A. L. P. Ribeiro, J. A. Behar, "Atrial fibrillation risk prediction from the 12-lead ECG using digital biomarkers and deep representation learning," European Heart Journal - Digital Health, 2021

Study of bio-markers from ECG to detect mortality

E. M. Lima*, A. H. Ribeiro*, G. M. M. Paixão et al., "Deep neural network estimated electrocardiographic-age as a mortality predictor," Nature Communications, vol. 12, 2021 *Equal contribution

Predicting risk of developing Atrial Fibrilation

S. Biton, S. Biton, A. H. Ribeiro, G. Miana, C. Moreira, A. L. P. Ribeiro, J. A. Behar, "Atrial fibrillation risk prediction from the 12-lead ECG using digital biomarkers and deep representation learning," European Heart Journal - Digital Health, 2021

Screening Chagas Diasease.



Figure: Performance on predicting Chagas Disease.

Study of bio-markers from ECG to detect mortality

E. M. Lima*, A. H. Ribeiro*, G. M. M. Paixão et al., "Deep neural network estimated electrocardiographic-age as a mortality predictor," Nature Communications, vol. 12, 2021 *Equal contribution

Predicting risk of developing Atrial Fibrilation

S. Biton, S. Biton, A. H. Ribeiro, G. Miana, C. Moreira, A. L. P. Ribeiro, J. A. Behar, "Atrial fibrillation risk prediction from the 12-lead ECG using digital biomarkers and deep representation learning," European Heart Journal - Digital Health, 2021

Screening Chagas Diasease.

Predicting Electrolyte values.



Figure: Performance on predicting Chagas Disease.

Outline

Automatic analysis of the electrocardiogram

Robustness and overparametrization

Research vision

Neural networks in critical applications



Figure: Effect of adversarial examples on ECG Classification.

Source: Han, X., Hu, Y., Foschini, L. et al. Deep learning models for electrocardiograms are susceptible to adversarial attack. Nature Medicine 26, 360–363 (2020).

Overparameterized models

Double descent

M. Belkin, D. Hsu, S. Ma, and S. Mandal , "Reconciling modern machine-learning practice and the classical bias-variance trade-off," PNAS (2019)

Overparameterized models

Double descent

M. Belkin, D. Hsu, S. Ma, and S. Mandal, "Reconciling modern machine-learning practice and the classical bias-variance trade-off," PNAS (2019)

Example:



Figure: nonlinear ARX mean squared error (MSE).

A. H. Ribeiro, J. N. Hendriks, A. G. Wills, T. B. Schön. "Beyond Occam's Razor in System Identification: Double-Descent when Modeling Dynamics". IFAC SYSID 2021. Honorable mention: Young author award

Overparameterized models

Double descent

M. Belkin, D. Hsu, S. Ma, and S. Mandal , "Reconciling modern machine-learning practice and the classical bias-variance trade-off," PNAS (2019)

Example:



Figure: nonlinear ARX mean squared error (MSE).

A. H. Ribeiro, J. N. Hendriks, A. G. Wills, T. B. Schön. "Beyond Occam's Razor in System Identification: Double-Descent when Modeling Dynamics". IFAC SYSID 2021. Honorable mention: Young author award

What is the role of high-dimensionality in model robustness?

What is the role of high-dimensionality in model robustness?

A. H. Ribeiro and T. B. Schön, "Overparameterized Linear Regression under Adversarial Attacks," arXiv:2204.06274. 2022.

$$y_i = \mathbf{x}_i^{\mathsf{T}}\boldsymbol{\beta} + \epsilon, i = 1, \cdots n$$

$$y_i = \mathbf{x}_i^{\mathsf{T}}\boldsymbol{\beta} + \epsilon, i = 1, \cdots n$$



$$y_i = \mathbf{x}_i^{\mathsf{T}}\boldsymbol{\beta} + \epsilon, i = 1, \cdots n$$

• estimate
$$\beta$$

• $\|\widehat{\beta}\|_2$ decays with $\frac{1}{\sqrt{\# \text{ features}}}$.



$$y_i = \mathbf{x}_i^{\mathsf{T}}\boldsymbol{\beta} + \epsilon, i = 1, \cdots n$$



• error =
$$y - x^{\mathsf{T}} \widehat{\beta}$$

- error = $y x^{\mathsf{T}}\hat{\beta}$
- lnput in the presence of an adv.: $x + \Delta x$

- error = $y x^{\mathsf{T}}\widehat{\beta}$
- lnput in the presence of an adv.: $x + \Delta x$
- ▶ ℓ_2 adversary *vs* ℓ_∞ adversary.

$$\begin{array}{c|c} & & \\ & \\ \ell_2 & \ell_{\infty} \end{array}$$

• error = $y - x^{\mathsf{T}}\hat{\beta}$

lnput in the presence of an adv.: $x + \Delta x$

▶ ℓ_2 adversary *vs* ℓ_∞ adversary.



 Understandment of overparameterized models brings them a step closer to its deployment in critical application.

- Understandment of overparameterized models brings them a step closer to its deployment in critical application.
- Linear regression: simplest case where both robustness or brittleness can be observed.

- Understandment of overparameterized models brings them a step closer to its deployment in critical application.
- Linear regression: simplest case where both robustness or brittleness can be observed.
- Adversarial training:

A. H. Ribeiro, D. Zachariah, and T. B. Schön, "Surprises in adversarially-trained linear regression," arXiv:2205.12695, May 2022.

Outline

Automatic analysis of the electrocardiogram

Robustness and overparametrization

Research vision

Build reliable and robust tools that can assist doctors in providing better patient care.

Theory: overparametrization and robustness **Methods:** learning from signals and systems **Applications:** diagnosis and prognosis of cardiac diseases



Building a research group

Students with applied problems but with theoretical and methodological challenges.

My current students:

 Daniel Gedon (Ph.D.) Unsupervised methods for ECG analysis

- Theogene Habineza (M.Sc.) AFib risk prediction
- Oscar Larson (M.Sc.) Adversarial attacks on the powergrid
- Philipp von Bachmann (M.Sc.) Electrolyte concentration prediction

Why I applied to Aalto University

- Top research and education institution
- Ample opportunity to collaboration:
 - Harri Lähdesmäki, Aki Vehtari, Simo Särkkä, Arno Solin, Leo Kärkkäinen
- Interdisciplinary research
 - Department of Computer Science
 - Department of Neuroscience and Biomedical Engineering
 - Department of Electrical Engineering and Automation
- Finnish Center for Artificial Intelligence - AI for health
- Good funding opportunities





My research philosophy

Research should be open, freely available and reproducible

- Contributions SciPy
- Open models and datasets for ECG analysis
 - CODE dset (n = 1.6M)
 - CODE-15% dataset (n > 200 000)
 - SaMI-Trop, CODE-test
 - International network collaboration

