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 AI-Health

Aalto University, Finland (Online)
June 2022

My background: Learning nonlinear dynamical systems

- **Example:** modeling deepwater oil well

Figure: Offshore extraction

Source: L. A. Aguirre, et al. "Development of soft sensors for permanent downhole Gauges in deepwater oil wells," Control Engineering Practice, 2017.

My background: Learning nonlinear dynamical systems

- **Example:** modeling deepwater oil well

y_t ! Pressure at the bottom;

Figure: Offshore extraction

Source: L. A. Aguirre, et al. "Development of soft sensors for permanent downhole Gauges in deepwater oil wells," Control Engineering Practice, 2017.

My background: Learning nonlinear dynamical systems

- **Example:** modeling deepwater oil well

y_t ! Pressure at the bottom;
u_t ! Pressure & temperature at platform;

Figure: Offshore extraction

Source: L. A. Aguirre, et al. "Development of soft sensors for permanent downhole Gauges in deepwater oil wells," Control Engineering Practice, 2017.

My background: Learning nonlinear dynamical systems

I **Example:** modeling deepwater oil well

y_t ! Pressure at the bottom;
u_t ! Pressure & temperature at platform;
t ! time;

Figure: Offshore extraction

Source: L. A. Aguirre, et al. “Development of soft sensors for permanent downhole Gauges in deepwater oil wells,” Control Engineering Practice, 2017.

My background: Learning nonlinear dynamical systems

- **Example:** modeling deepwater oil well

y_t ! Pressure at the bottom;
u_t ! Pressure & temperature at platform;
t ! time;

- Auto-regressive model:

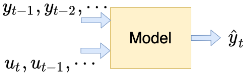


Figure: Offshore extraction

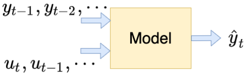
Source: L. A. Aguirre, et al. "Development of soft sensors for permanent downhole Gauges in deepwater oil wells," Control Engineering Practice, 2017.

My background: Learning nonlinear dynamical systems

- **Example:** modeling deepwater oil well

  $y_t \to$ Pressure at the bottom;
  $u_t \to$ Pressure & temperature at platform;
  $t \to$ time;

- Auto-regressive model:



- Data-driven methods:

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

Figure: Offshore extraction

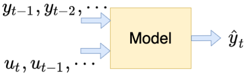
Source: L. A. Aguirre, et al. "Development of soft sensors for permanent downhole Gauges in deepwater oil wells," Control Engineering Practice, 2017.

My background: Learning nonlinear dynamical systems

- **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:

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

- Long-term predictions:

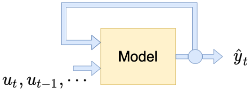


Figure: Offshore extraction

Source: L. A. Aguirre, et al. "Development of soft sensors for permanent downhole Gauges in deepwater oil wells," Control Engineering Practice, 2017.

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 tradeo s between extra- exibility 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.
- 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.
- 18 million deaths in 2019 (32%).
- Myocardial infarction 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...

Left: ECG signal **Right:** Electrode placement.

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

Telehealth and automatic tools

- The Telehealth Center of Minas Gerais

Telehealth and automatic tools

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

Telehealth and automatic tools

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

Telehealth and automatic tools

- 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

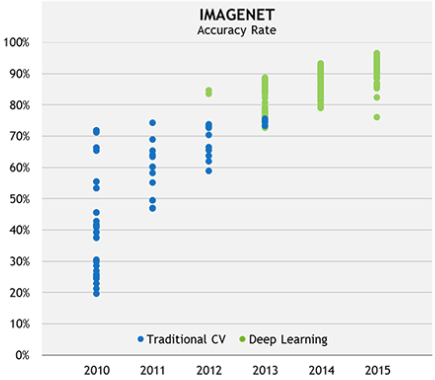


Figure: Models accuracy on ImageNet benchmark. In 2012, LeNet-5 wins ILSVRC.

Deep learning for automatic diagnosis

- CODE dataset: historical data 2010 to 2017.
  - 2.3 M ECGs from n =1.6M 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 =1.6M patients
  - 6 abnormalites
- CODE-test: annotate by 3 cardiologists (n=827).
- Develop and evaluate deep neural network

A. H. Ribeiro , M.H. Ribeiro, Paixao, 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. ! 4th year cardiology residents
emerg. ! 3rd year emergency residents
stud. ! 5th year Medical students

Results

Bootstrapped F1 score values

	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. ! 4th year cardiology residents

emerg. ! 3rd year emergency residents

stud. ! 5th year Medical students

Discussion

- Proof of concept study.

Discussion

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

Discussion

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

Discussion

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

D. Gedon et al. , S. Gustafsson , E. Lampa, A. H. Ribeiro , M. J. Holzmann, T. B. Schön, J. Sundstrom, "ResNet-based ECG Diagnosis of Myocardial Infarction in the Emergency Department," Workshop at NeurIPS, 2021. Under review

Myocardial Infarction

Can we detect non-STEMI cases from the ECG?

Myocardial Infarction

Can we detect non-STEMI cases from the ECG?

Diagnosis:

- ST-elevation MI (STEMI)! ECG.
- non-STEMI! blood testing

Left: No MI. Right: STEMI.

D. Gedon et al. , S. Gustafsson , E. Lampa, A. H. Ribeiro , M. J. Holzmann, T. B. Schön, J. Sundstrom, "ResNet-based ECG Diagnosis of Myocardial Infarction in the Emergency Department," Workshop at NeurIPS, 2021. Under review

Myocardial Infarction

Can we detect non-STEMI cases from the ECG?

Diagnosis:

- ST-elevation MI (STEMI)! ECG.
- non-STEMI ! blood testing

Dataset:

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

Left: No MI. Right: STEMI.

D. Gedon et al. , S. Gustafsson , E. Lampa, A. H. Ribeiro , M. J. Holzmann, T. B. Schön, J. Sundstrom, "ResNet-based ECG Diagnosis of Myocardial Infarction in the Emergency Department," Workshop at NeurIPS, 2021. Under review

Myocardial Infarction

Can we detect non-STEMI cases from the ECG?

Diagnosis:

- ST-elevation MI (STEMI)! ECG.
- non-STEMI ! blood testing

Dataset:

- Stockholm emergency: 2007 to 2016
- 400 000 ECGs
- NSTEMI 1%. STEMI 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).

D. Gedon et al. , S. Gustafsson , E. Lampa, A. H. Ribeiro , M. J. Holzmann, T. B. Schön, J. Sundstrom, "ResNet-based ECG Diagnosis of Myocardial Infarction in the Emergency Department," Workshop at NeurIPS, 2021. Under review

Data-driven methods to detect new conditions

- 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

Data-driven methods to detect new conditions

- 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 brillation risk prediction from the 12-lead ECG using digital biomarkers and deep representation learning," European Heart Journal - Digital Health, 2021

Data-driven methods to detect new conditions

- 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 brillation 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.

Data-driven methods to detect new conditions

- 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 brillation 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: E ect of adversarial examples on ECG Classi cation.

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-o ," 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-o ," 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 Identi cation: 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-o ," 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 Identi cation: 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.

Overparameterized models and robustness

Dataset $(x_i; y_i)$, $\dim(x_i) =\ m$:

$$y_i = x_i^T \quad + \quad ; i = 1; \qquad n$$

Overparameterized models and robustness

Dataset $(x_i; y_i)$, $\dim(x_i) = m$:

$$y_i = x_i^T \quad + \quad ; i = 1; \quad n$$

- estimate b

Overparameterized models and robustness

Dataset $(x_i; y_i)$, $\dim(x_i) = m$:

$$y_i = x_i^T \quad + \quad ; i = 1; \quad n$$

- estimate b
- $\| b \|_2$ decays with $\sqrt{\frac{1}{\#\ \text{features}}}$.

Overparameterized models and robustness

Dataset $(x_i ; y_i)$, $\dim(x_i) = m$:

$$y_i = x_i^{T} \quad + \quad ; i = 1; \quad n$$

- estimate b
- $k\,b\,k_2$ decays with $p \frac{1}{\#\text{ features}}$.
- $k\,b\,k_1$ goes to constant.

Left: `$_2$-norm. Right: `$_1$-norm.

Adversarial attacks

- error = $y - x^T b$

Adversarial attacks

- error = $y \quad x^T b$
- Input in the presence of an adv$x$: + $\quad x$

Adversarial attacks

- error = $y \quad x^T b$
- Input in the presence of an adv: $x + \quad x$
- ℓ_2 adversary vs ℓ_1 adversary.

ℓ_2 ℓ_1

Adversarial attacks

- error = $y \quad x^T b$
- Input in the presence of an adv.: $x + \quad x$
- ℓ_2 adversary vs ℓ_1 adversary.

$$\ell_2 \qquad \ell_1$$

- $E\{\text{error}^2\} + {}^2\|b\|^2 \quad E\{(\text{adv. error})^2\} \quad E\{\text{error}\} + \|b\|^{2}$

Discussion

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

Discussion

- 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.

Discussion

- 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

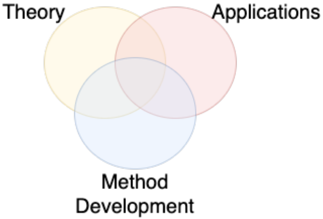
Robustness and overparametrization

Research vision

Research goals

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 ($n = 1.6$M)
  - CODE-15% dataset ($n > 200\,000$)
  - SaMI-Trop, CODE-test
  - International network collaboration