AI Application in Health

Javier Serrano
Applied Machine Learning
Master in Data Science and Analytics

Strathmore University
@iLabAfrica Centre

Sumary

Machines are beating us at everything

"They, the machines, will make decisions for us, they will govern us."
Ethics for Machines, Jose Ignacio Latorre

- Initially, machines replaced us in physical labor.
- Then in calculation.
- Now there are machines making decisions that are less and less trivial.
- In the future, machines will increasingly make ethical decisions.

And we like to **delegate decisions**.

How do Artificial Intelligences Learn?

Two strategies:

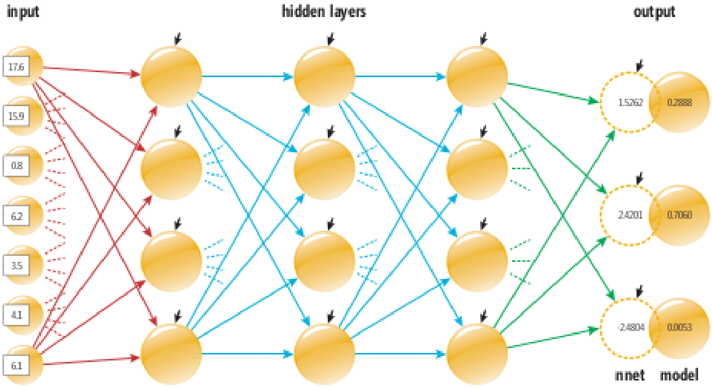
Expert systems (80's):

- Humans program their knowledge into detailed algorithms.
- Advanced statistics.
- Bayesian inference.
- Ex: Medical diagnosis.
- It was a **failure**.
- **AI winter**.

Neural networks:

- Learning from examples.
- Copy the human brain, not its logic.
- Learning = Training.

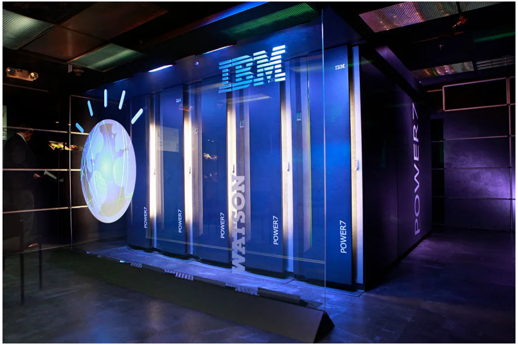
Neural networks learn from examples



Great successes: AlphaGo

Medical Assistant: IBM Watson

Delegation of Decisions



- Doctors can now have a virtual assistant.
- Trained with thousands of cases.
- Thousands of treatments.
- Detailed casuistry. /item Inference from inaccurate data.

Is this the way?

Sumary

What are the main problems in Healthcare Systems?

(with people living 100 years)

In developed countries

The healthcare expenditures are growing in an unsustainable way

In underdeveloped countries

There are not enough resources for taking care of a growing population

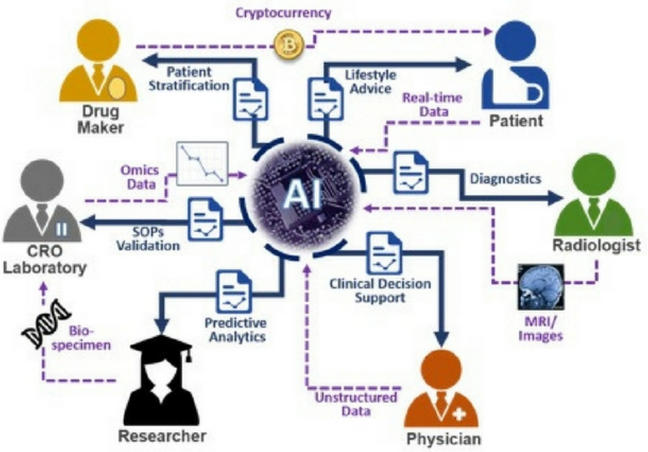
How can AI help to solve this problems?

Areas where AI can Help

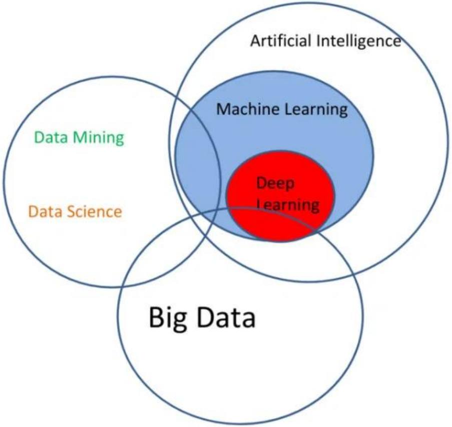
AI can help to reduce costs in two way (at least)

1. **Clinical Decision Making**
   - Avoiding misdiagnosis and unnecessary treatments
   - Personal recommend treatments
   - Preventive medicine. Models for risk prediction
2. Operational planning. Resource allocation

Artificial Intelligence and Medicine



Differences between AI, Machine learning and Data Analytics

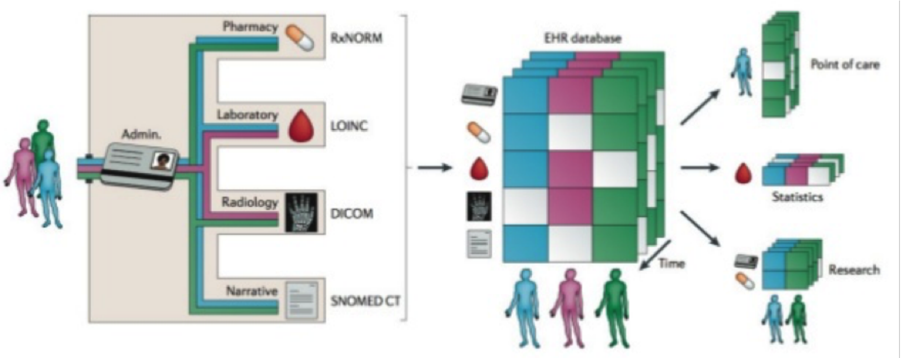


Why Deep Learning Now?

From 2012 ....

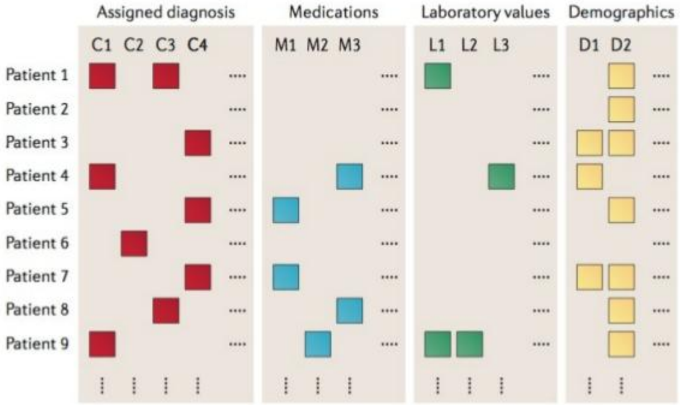
- Big Datasets
  - Images
  - Vital signals/Sensors (ICU, wearables, ...)
  - EHR
  - Laboratory Data
  - Genome sequences
  - Events sequences
- GPU availability (High Performance Computing)
- Cloud Computing
- Open source algorithms (Tensorflow, Pandas, SkLearn, ...)

Medical Data & Electronic Health Record (EHR)



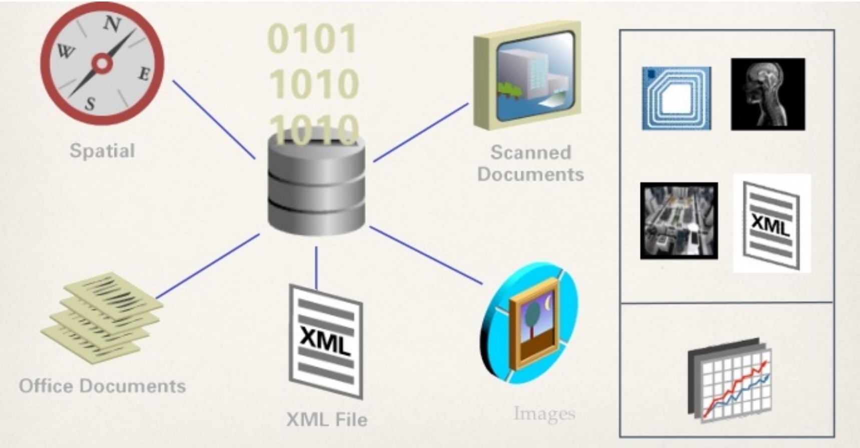
What Kind of Clinical Data Can Be Found in a Hospital?

Structured Data - The Electronic Health Register (EHR)

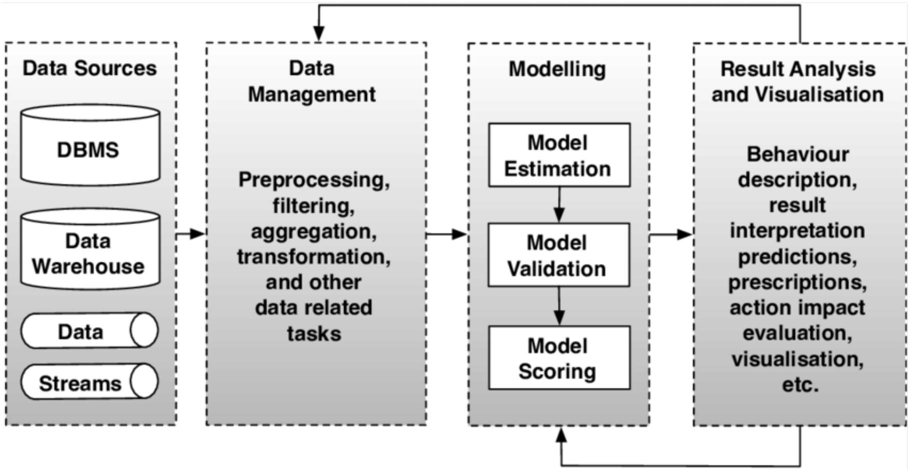


What Kind of Clinical Data Can Be Found in a Hospital?

Unstructured Data: Images, Clinical reports, Time Series

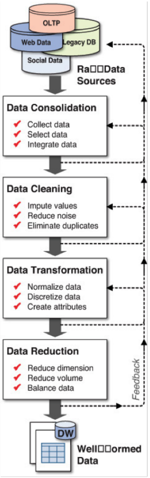


Typical Workflow in Data Analytics

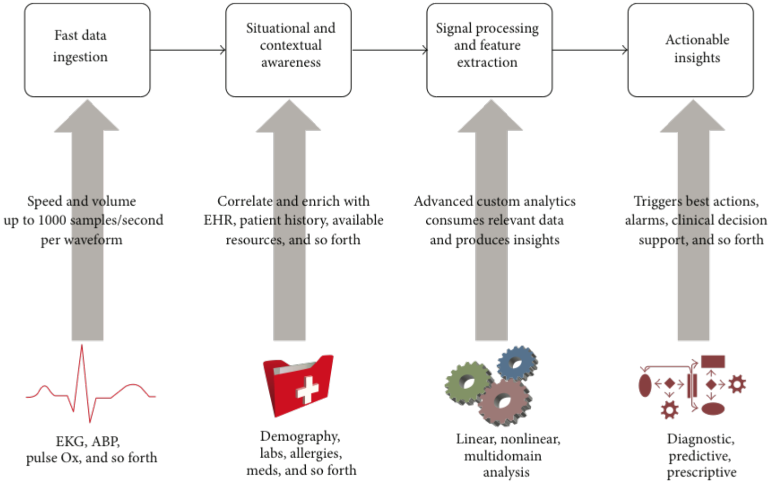


Data Preprocessing

- Time consuming
- Very important
- Anonymization is a previous step to preprocessing



Big data Processing



Sumary

Application to Physiological Signals: Arousals detection

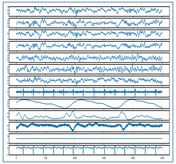
Knowledge extraction based on wavelets and DNN for physiological signals (PS): Arousals case

- **Goal:** Detection of arousal zones when sleep
- Useful to detect sleep disorders
- Sleep (-4-9 hours) = ... non-REM + REM + non-REM + ... (REM: Rapid Eye Movement)
- Signal we analyze: ECG + VOmax + EEG + EOM
- Non-balanced data
- Medical collaboration is needed.

Application to Physiological Signals: Arousals detection

Knowledge extraction based on wavelets and DNN for physiological signals (PS): Arousals case

1. Slide windows, of 13 PS



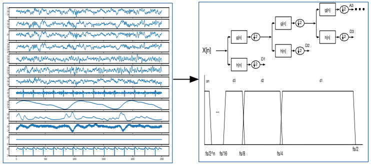
Medical decision support based on discrete wavelet transform for feature extraction and classification of 13 PS, >1850 patients

Application to Physiological Signals: Arousals detection

Knowledge extraction based on wavelets and DNN for physiological signals (PS): Arousals case

1. Slide windows, of 13 PS

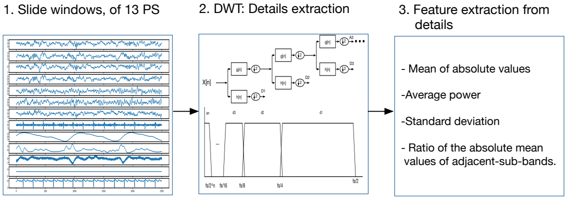
2. DWT: Details extraction



Medical decision support based on discrete wavelet transform for feature extraction and classification of 13 PS, >1850 patients, 6M registers.

Application to Physiological Signals: Arousals detection

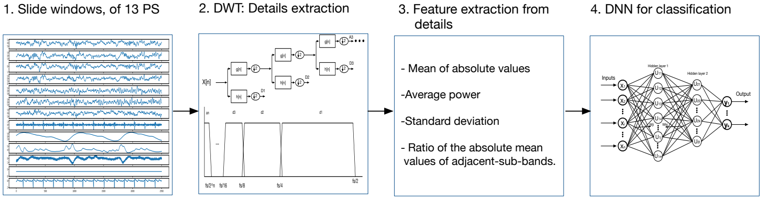
Knowledge extraction based on wavelets and DNN for physiological signals (PS): Arousals case



Medical decision support based on discrete wavelet transform for feature extraction and classification of 13 PS, >1850 patients, 6M registers. Information reduction from 32500 samples per patient to 312

Application to Physiological Signals: Arousals detection

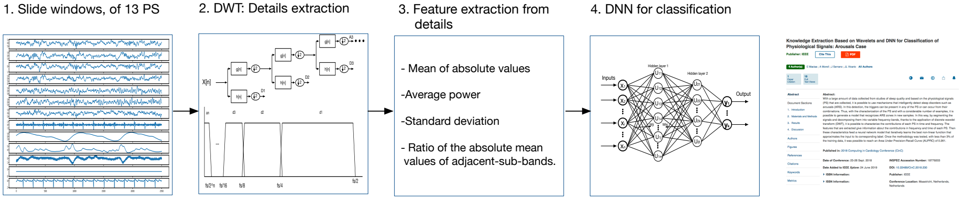
Knowledge extraction based on wavelets and DNN for physiological signals (PS): Arousals case



Medical decision support based on discrete wavelet transform for feature extraction and classification of 13 PS, >1850 patients, 6M registers. Information reduction from 32500 samples per patient to 312

Application to Physiological Signals: Arousals detection

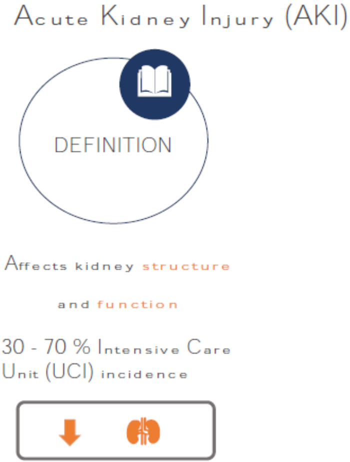
Knowledge extraction based on wavelets and DNN for physiological signals (PS): Arousals case



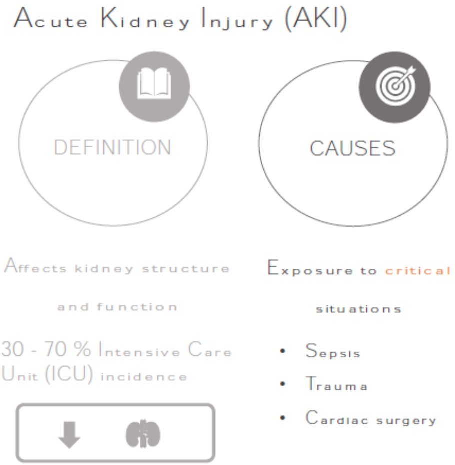
Medical decision support based on discrete wavelet transform for feature extraction and classification of 13 PS, >1850 patients, 6M registers. Information reduction from 32500 samples per patient to 312

Sumary

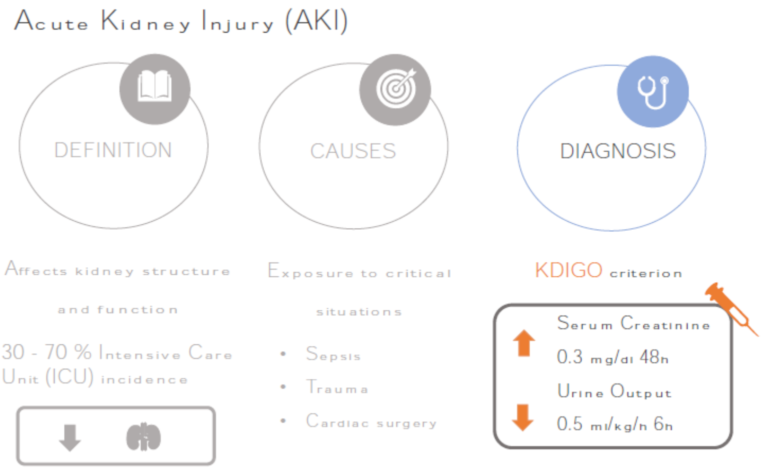
Case 1: Mortality Prediction for AKI



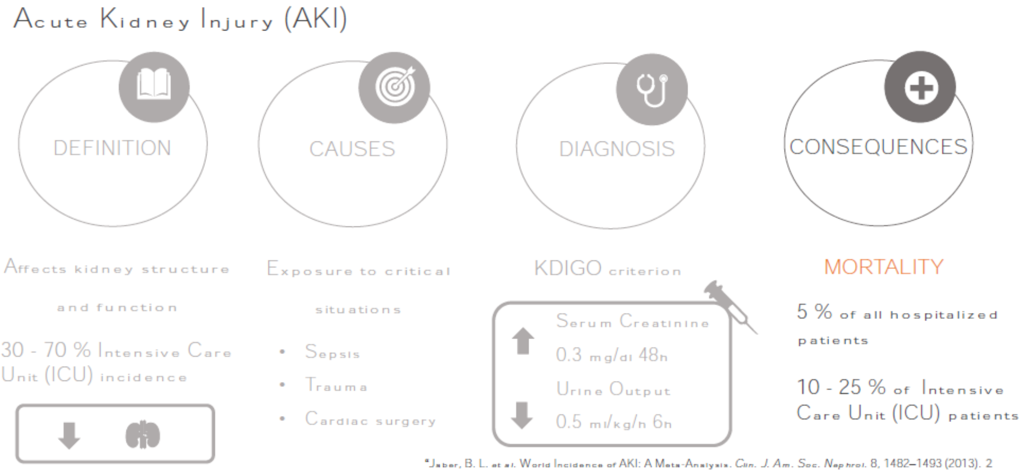
Case 1: Mortality Prediction for AKI



Case 1: Mortality Prediction for AKI



Case 1: Mortality Prediction for AKI



Case 1: Mortality Prediction for AKI

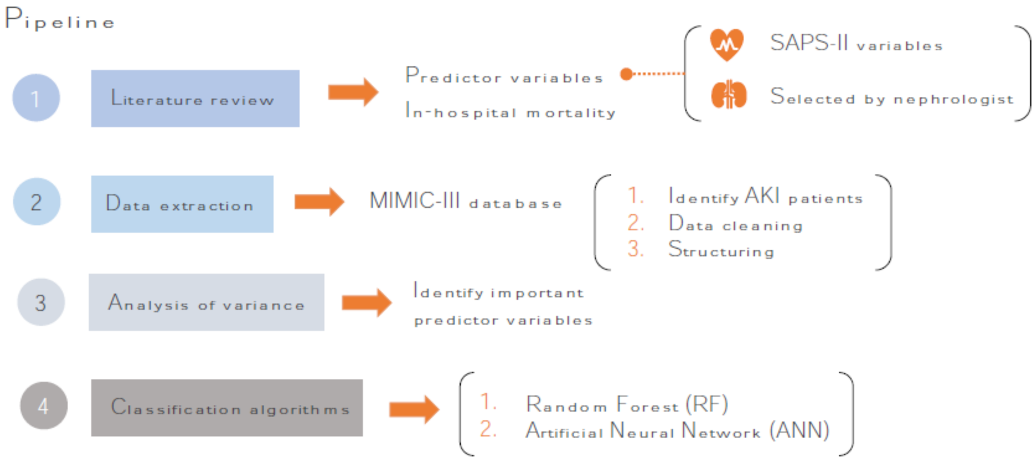
Objectives

General **Predict** mortality in-hospital in patients with AKI in the ICU

Secondary Select and analyze the most important predictors variables (**Knowledge Discovery**)

Case 1: Mortality Prediction for AKI

Methods



Methods

Variables and Metrics

Variables

Factors	SAPS-II	Selected by an expert
Age	Heart Rate	Creatinine
Type of admission	Systolic Blood Pressure	Glucose
Relevant diseases	Temperature	Platelet Count
Gender	GCS	Red Blood Cells
Number of admissions	PaO_2	Chloride
	FiO_2	Anion Gap
	Urine Output	Magnesium
	BUN	PT
	Sodium	PTT
	Potassium	Phosphate
	Bicarbonate	Calcium
	Bilirubin	pH
	WBC	pCO_2
		pO_2
		Lactate
		Lymphocytes
		Eosinophils
		Neutrophils
		Alanine Aminotransferase
		Aspartate Aminotransferase
		Alkaline Phosphatase

Metrics

Confusion Matrix		Predicted	
		Alive	Dead
Actual	Alive	TN	FP
	Dead	FN	TP

General performance:

- Area Under Receiving Operation Characteristic (AUC)
- Accuracy $= \frac{TP+TN}{TP+TN+FP+FN}$

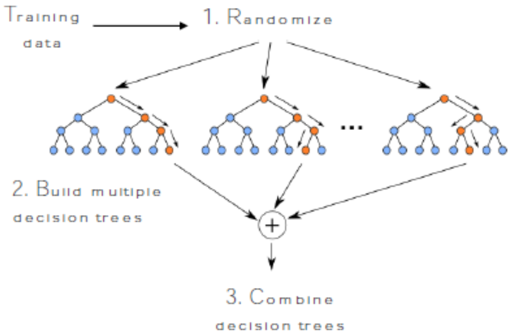
$\left(\text{Biased parameter for unbalanced groups} \right)$

Classification Algorithms

Random Forest

Some features

- RF is very easy to set up. Good algorimth for starting
- Used as baseline system
- Even with small data works properly
- Rate the importance of the attributes

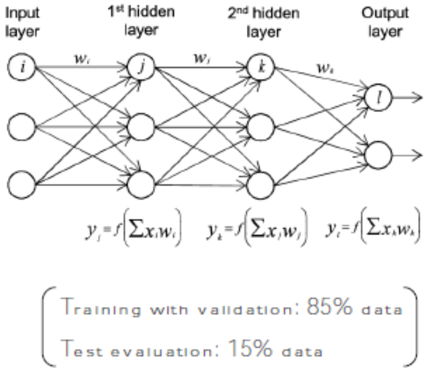


Classification Algorithms

Artificial Neural Networks

Some features

- Usually gets the best performance
- Lots of parameter to tune, difficult
- With small data tends to overfit or do not get a good performance



Results

Comparison performance: Clinically valuable metrics

	Sensitivity		NPV		AUROC	
	RF	ANN	RF	ANN	RF	ANN
1. SAPS-II 4153	63.7%	68.8%	78.7%	80.3%	0.79	0.77
2. SAPS-II stat. 6754	89.3%	81.5%	92.4%	89.0%	0.77	0.76
3. SAPS-II stat. 15435	74.7%	77.7%	89.9%	90.6%	0.82	0.82
4. SAPS-II 15435	87.1%	85.7%	94.2%	93.8%	0.87	0.87
5. SAPS-II + analytics 15435	76.5%	79.9%	91.7%	92.5%	0.85	0.85

Results

Comparison performance: Other metrics

	Specificity		Precision		Accuracy	
	RF	ANN	RF	ANN	RF	ANN
1. SAPS-II 4153	80.7%	76.6%	66.5%	63.9%	74.3%	73.6%
2. SAPS-II stat. 6754	49.9%	57.2%	40.6%	42.2%	60.8%	63.9%
3. SAPS-II stat. 15435	74.1%	71.1%	48.8%	46.9%	74.3%	72.7%
4. SAPS-II 15435	69.8%	71.4%	48.7%	49.7%	74.1%	75.0%
5. SAPS-II + analytics 15435	79.6%	75.9%	53.5%	50.5%	78.9%	76.9%

Discussion

Strengths and Drawbacks

Strengths

- Good performance in relevant metrics
- High potentials on clinical applications
- Use of a public-available and large-scale database

Drawbacks

- Insufficient results on metrics related to resource management
- Retrospective study
- Not able to implement yet in clinics

Case 1: Mortality Prediction for AKI

Conclusions

Predictive model of in-hospital mortality for AKI patients in the ICU

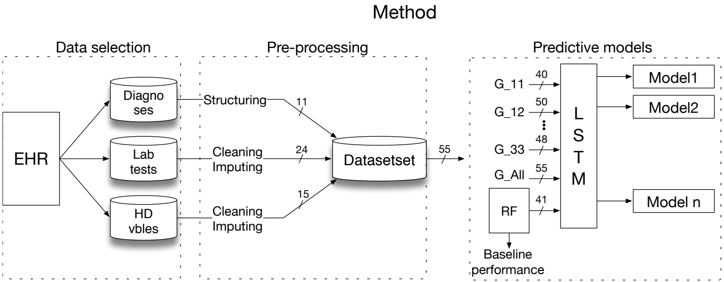
- Increase on the number of subjects
- Increase on the number of predictor variables

High potential in clinical practice

- Take decisions in critical situations
- Redistribute/allocate resources in ICU (**Ethical Issues**)

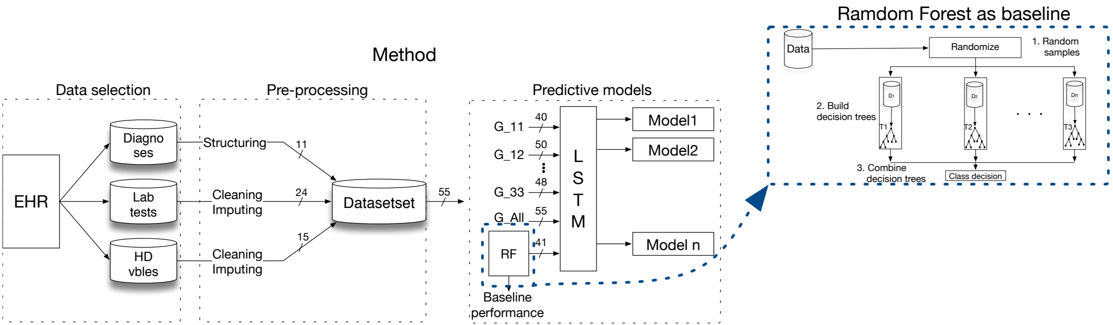
Case 2: Mortality Prediction in ESRD

Mortality prediction enhancement in end-stage renal disease: A machine learning approach



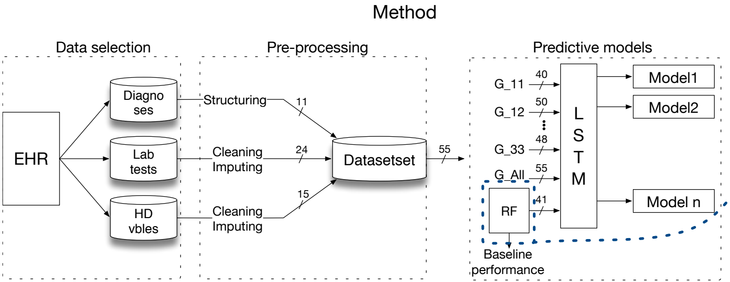
Case 2: Mortality Prediction in ESRD

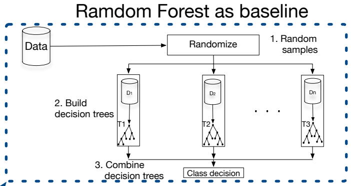
Mortality prediction enhancement in end-stage renal disease: A machine learning approach

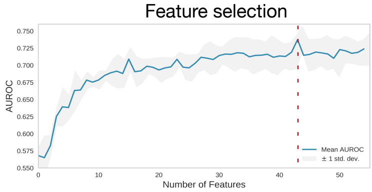


Case 2: Mortality Prediction in ESRD

Mortality prediction enhancement in end-stage renal disease: A machine learning approach

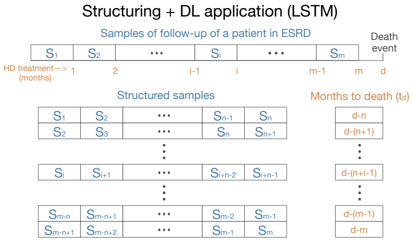






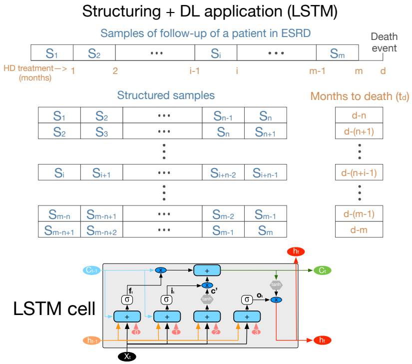
Case 2: Mortality Prediction in ESRD

Mortality prediction enhancement in end-stage renal disease: A machine learning approach



Case 2: Mortality Prediction in ESRD

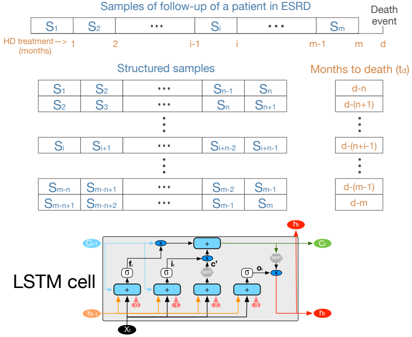
Mortality prediction enhancement in end-stage renal disease: A machine learning approach

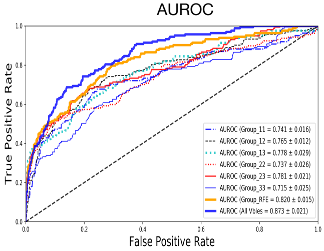


Case 2: Mortality Prediction in ESRD

Mortality prediction enhancement in end-stage renal disease: A machine learning approach

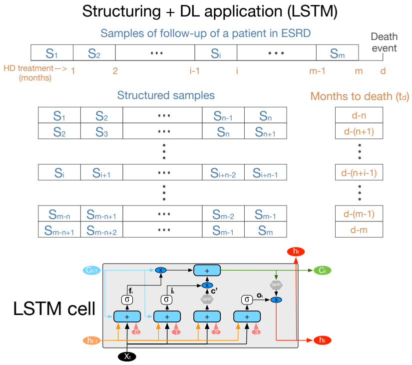
Structuring + DL application (LSTM)

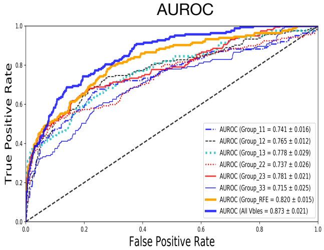




Case 2: Mortality Prediction in ESRD

Mortality prediction enhancement in end-stage renal disease: A machine learning approach





Download PDF Share Export

Informatics in Medicine Unlocked
Volume 19, 2020, 100351

Mortality prediction enhancement in end-stage renal disease: A machine learning approach

Edwar Macias, Antoni Morell, Javier Serrano, Jose Lopez Vicario, Jose Ibeas

Show more

https://doi.org/10.1016/j.imu.2020.100351
Under a Creative Commons license
Get rights and content
open access

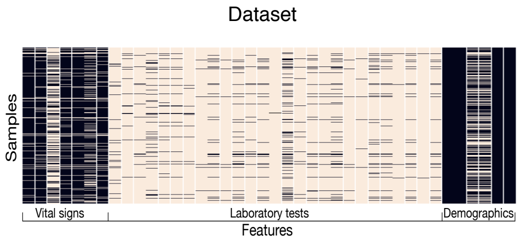
Highlights

- Machine learning approaches can lead to a paradigm shift in the analysis of predictive factors for mortality in ESRD.
- The massive use of variables together with artificial neural networks improves predictive models of mortality in ESRD.
- Machine learning approaches can reveal causal relationships in variable not explored before by the expert staff.

Sumary

Application to EHR + Physiological Signals

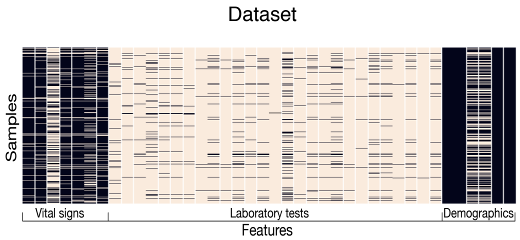
Novel Imputing Method and Deep Learning Techniques for Early Prediction of Sepsis in ICU



- Challenge: predict sepsis 6 hours before the clinical prediction of sepsis
- 3 ICUs
- +2000 patients

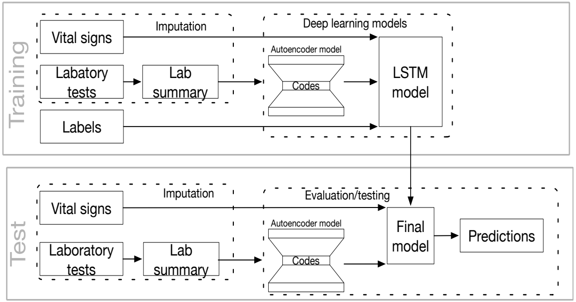
Application to EHR + Physiological Signals

Novel Imputing Method and Deep Learning Techniques for Early Prediction of Sepsis in ICU



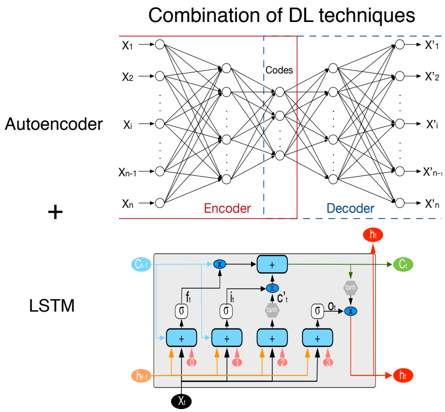
- Challenge: predict sepsis 6 hours before the clinical prediction of sepsis
- 3 ICUs
- +2000 patients

Proposed method: Combine DL techniques for imputation and prediction



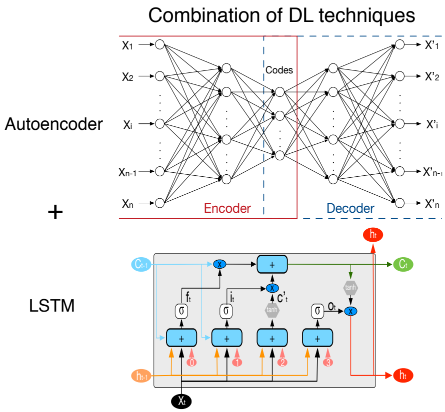
Application to EHR + Physiological Signals

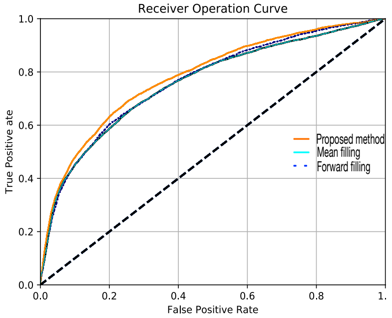
Novel Imputing Method and Deep Learning Techniques for Early Prediction of Sepsis in ICU



Application to EHR + Physiological Signals

Novel Imputing Method and Deep Learning Techniques for Early Prediction of Sepsis in ICU



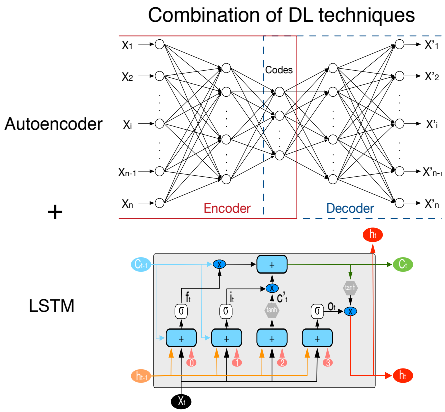


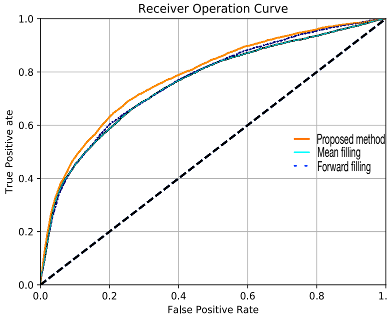
Results

- 38/78 in challenge
- 27% dimensionality reduction
- 10% better in Utility

Application to EHR + Physiological Signals

Novel Imputing Method and Deep Learning Techniques for Early Prediction of Sepsis in ICU





Novel Imputing Method and Deep Learning Techniques for Early Prediction of Sepsis in Intensive Care Units

Publisher: IEEE | Cite This | PDF

6 Author(s) E Macias ; G Boquet ; J Serrano ; JL Vicario ; J Ibeas ; A Morel All Authors

32 Full Text Views

Abstract

Document Sections
1. Introduction
2. Materials and methods
3. Results
4. Conclusion

Authors
Figures
References
Keywords
Metrics

Abstract:
It is possible to exploit the predictive capacity of data collected in intensive care units (ICU) with a high rati of missing values. Combining several sources of information, a considerable number of missing values are generated. In this manuscript, an alternative approach to impute this type of data, together with the use of deep learning techniques to improve the early detection of sepsis in ICU is proposed. Initially, laboratory tests are separated and summarized. Then, their most representative information is extracted by taking codes from an autoencoder. This information is combined with the rest of the variables and used to exploit temporal dependencies through long short-term memory recurrent neural networks. With the proposed approach our team, WIN-UAB, was ranked in the position 38/78 with a utility score (defined in the the PhysioNet/Computing in Cardiology Challenge 2019) of 0.241 on the full test set. The predictive capacity o the proposed solution demonstrated the potential of integrating an alternative approach for imputing variables with a high ratio of missing values. In terms of dimensionality reduction, it is possible to reduce 27% of features through the codes of autoencoders.

Published in: 2019 Computing in Cardiology (CinC)

Date of Conference: 8-11 Sept. 2019
INSPEC Accession Number: 19381149
Date Added to IEEE *Xplore*: 24 February 2020
DOI: 10.23919/CinC49843.2019.9005794
ISBN Information:
Publisher: IEEE
ISSN Information:
Conference Location: Singapore, Singapore, Singapore

Results

- 38/78 in challenge
- 27% dimensionality reduction
- 10% better in Utility

Sumary

Goals

Applied Machine Learning

- Understand the intuition behind modern machine learning methods as well as a bit more formal understanding of how, why, and when they work.
- Apply the machine learning techniques over data to solve practical problems.

Contents

Applied Machine Learning

1. Fundamentals of ML item Data Pre-processing
2. Learning from data
3. Supervised learning techniques
4. Non-supervised learning techniques
5. Artificial Neural Networks

Bibliography

Applied Machine Learning

1. *Hands on Machine Learning*, A. Geron. 2/3th edition
https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=
https://powerunit-ju.com/wp-content/uploads/2021/04/
Aurelien-Geron-Hands-On-Machine-Learning-with-Scikit-Learn-Keras-and-Te
-Concepts-Tools-and-Techniques-to-Build-Intelligent-Systems-OReilly-Med
pdf&ved=2ahUKEwiSoZb4j--HAxW7VUEAHceHHzkQFnoECEUQAQ&usg=
AOvVaw0-0QzVlPQS-U-V6EAvoh8G

VLE + Teams

Applied Machine Learning

The course is based on the Virtual Classroom.

- Communication
- Materials
- Deliveries
- Marks

Evaluation

Applied Machine Learning

- Sitting exam: 60%
- Assignments: 30%
- Capstone project: 10%