MESTRADO
ENGENHARIA ELETRÔNICA
E COMPUTAÇÃO

Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission

TÉCNICAS DE DATA MINING E MACHINE LEARNING APLICADAS ÀS CIÊNCIAS BIOMÉDICAS

Rodrigo Blanke Lambrecht
Pelotas, 2018

COLABORADORES

- Rich Caruana: Microsoft Research
- Paul Koch: Microsoft Research
- Yin Lou: LinkedIn Corporation
- Marc Sturm: NewYork-Presbyterian Hospital
- Johannes Gehrke: Microsoft
- Noémie Elhadad: Columbia University

PROPOSTA

- Aplicação de um método de aprendizagem baseado em **modelos aditivos generalizados de alto desempenho** (GA^{2}M) fazendo a implementação em dois estudos de caso:
  - **PNEUMONIA RISK:** Prever o risco de morte por pneumonia;;
  - **30-DAY READMISSION:** Readmissão hospitalar em 30 dias.

PNEUMONIA RISK

Prever a probabilidade de morte (DPO) de pacientes com Pneumonia.

- **14199** Pacientes
  - 10,86% dos pacientes(1542 pacientes) morreram de pneumonia
- **46** Atributos para cada paciente
- Conjunto de treino e testes: **70:30**
  - **9847** Pacientes para treino
  - **4352** Pacientes para teste

PNEUMONIA RISK - ATRIBUTOS

Patient-history findings			
chronic lung disease	-	age	C
re-admission to hospital	-	gender	-
admitted through ER	-	diabetes mellitus	-
admitted from nursing home	-	asthma	-
congestive heart failure	-	cancer	-
ischemic heart disease	-	number of diseases	C
cerebrovascular disease	-	history of seizures	-
chronic liver disease	-	renal failure	-
history of chest pain	-		
Physical examination findings			
diastolic blood pressure	C	wheezing	-
gastrointestinal bleeding	-	stridor	-
respiration rate	C	heart murmur	-
altered mental status	-	temperature	C
heart rate	C		

PNEUMONIA RISK - ATRIBUTOS

Laboratory findings			
liver function tests	-	BUN level	C
glucose level	C	creatinine level	C
potassium level	C	albumin level	C
hematocrit	C	WBC count	C
percentage bands	C	pH	C
pO2	C	pCO2	C
sodium level	C		
Chest X-ray findings			
positive chest x-ray	-	lung infiltrate	-
pleural effusion	-	pneumothorax	-
cavitation/empyema	-	chest mass	-
lobe or lung collapse	-		

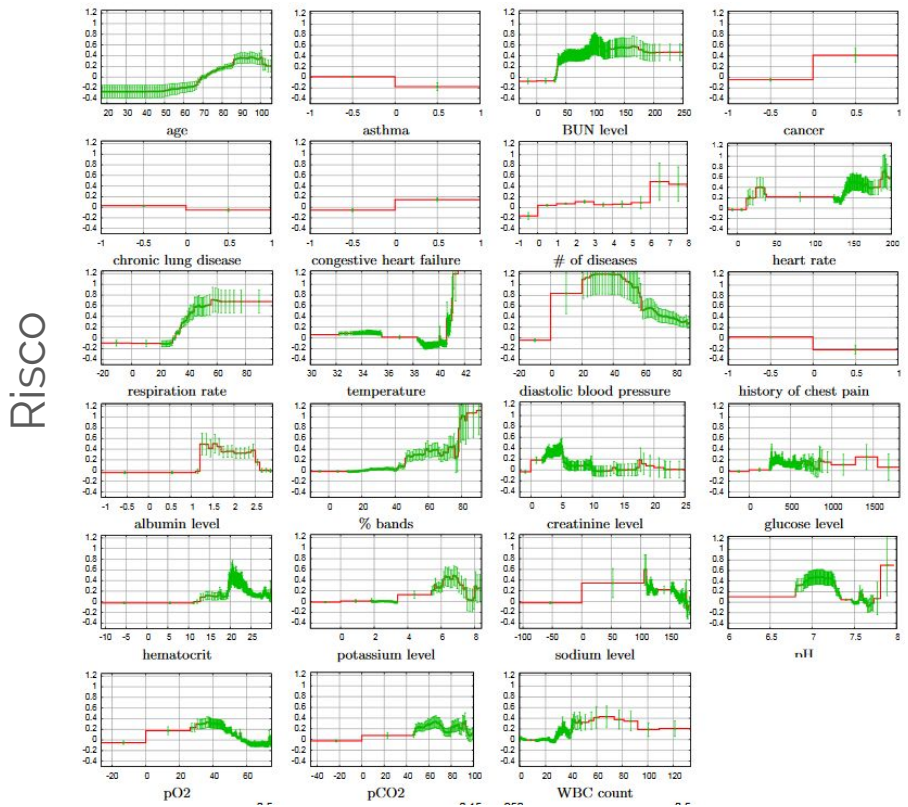
PNEUMONIA RISK

Os modelos GAM/GA^{2}M são treinados nesses dados usando 100 iterações de bagging (ensacamento):

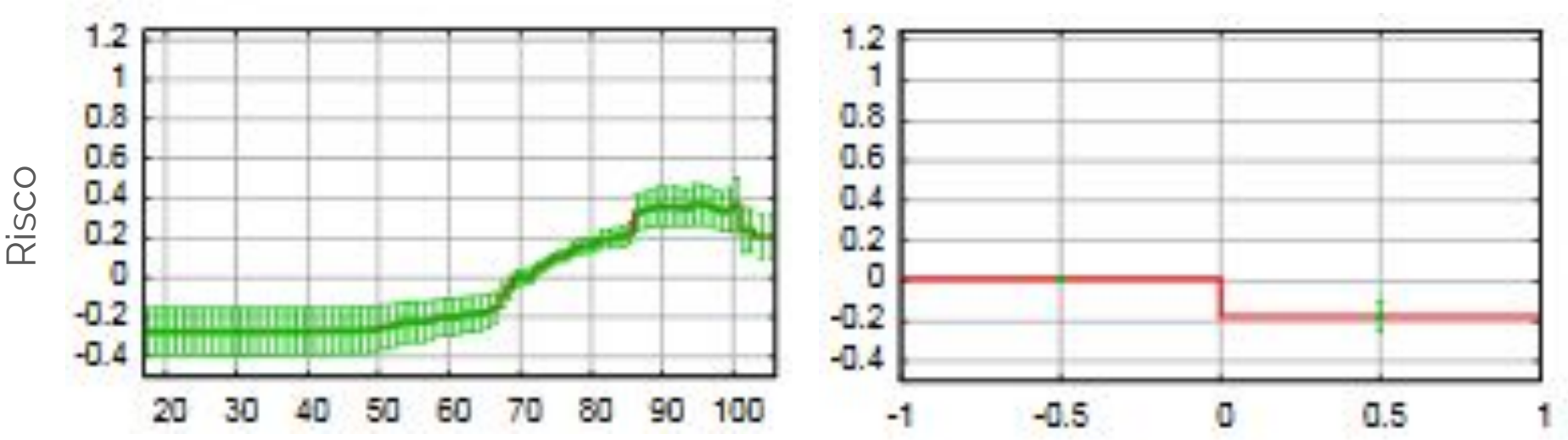
- Reduzir o overfitting
- Melhorar a precisão

PNEUMONIA RISK

Gráfico com score de risco para alguns atributos utilizados nas iterações de bagging;



PNEUMONIA RISK



PNEUMONIA RISK

Model	Pneumonia	Readmission
Logistic Regression	0.8432	0.7523
GAM	0.8542	0.7795
GA^2M	0.8576	0.7833
Random Forests	0.8460	0.7671
LogitBoost	0.8493	0.7835

30-DAY READMISSION

Prever a necessidade de retorno ao Hospital após 30 dias de liberados.

- **296724** Pacientes
  - 8,91% dos pacientes são readmitidos em 30 dias
- **3956** Atributos para cada paciente
- Conjunto de treino e testes: **66:34**
  - **195901** Pacientes para treino
  - **100823** Pacientes para teste

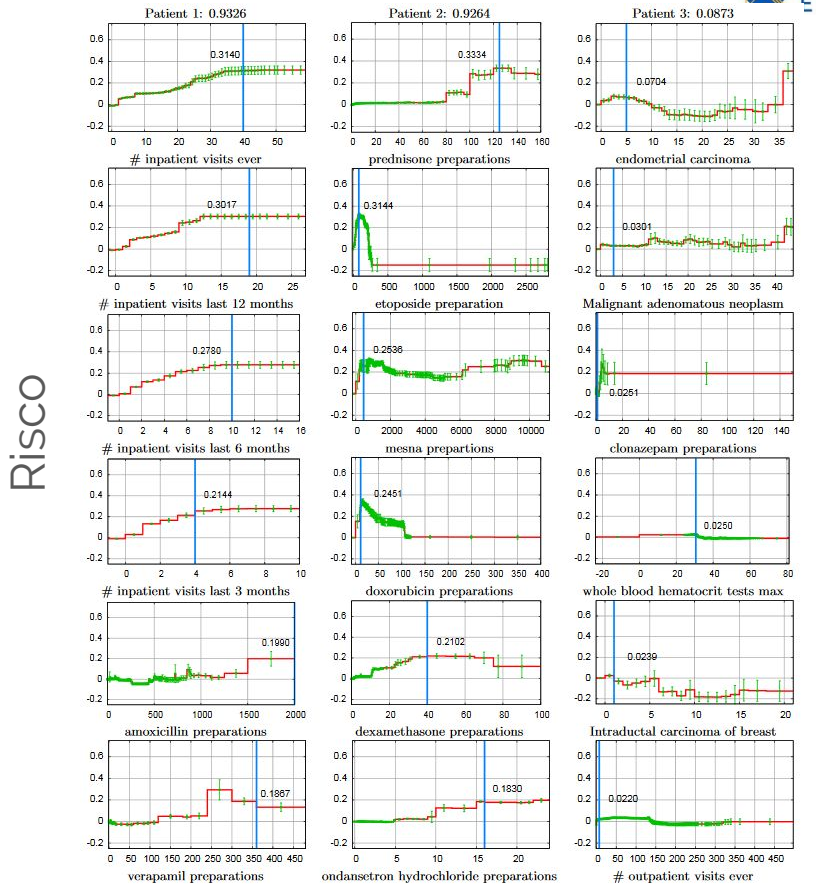
30-DAY READMISSION

Foi utilizado 10 iterações de bagging

- Para treinar 10 modelos levaria de 2 a 3 dias em um grupo de 10 computadores de uso geral;
- Portanto examinaram as previsões para 3 pacientes:
    - 2 com probabilidade de readmissão muito alta e 1 com probabilidade típica.

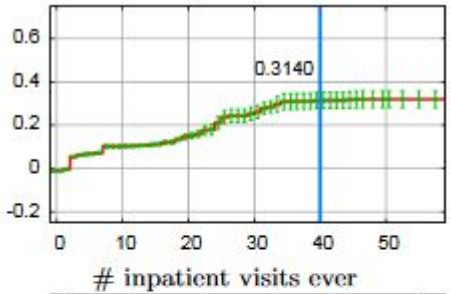
PNEUMONIA RISK

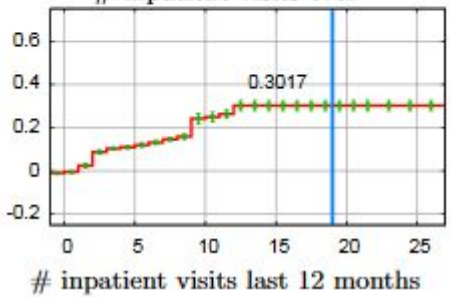
Gráfico com score de risco para alguns atributos dos 3 pacientes

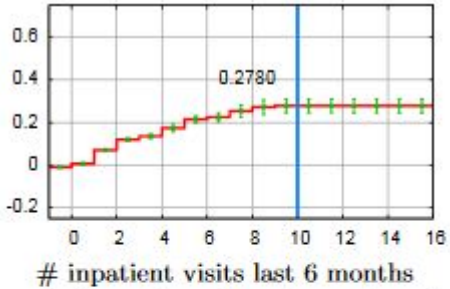


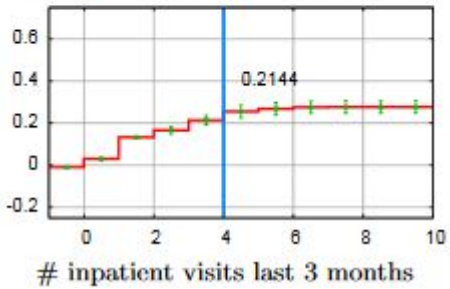
PNEUMONIA RISK

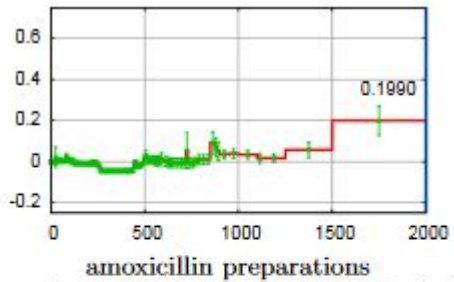
Paciente 1:

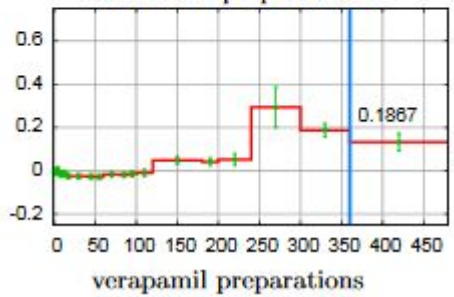






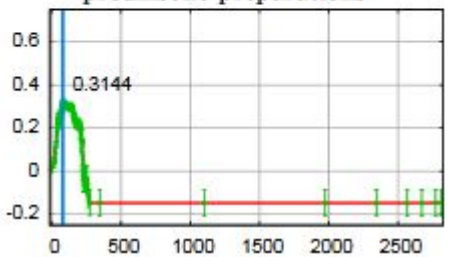


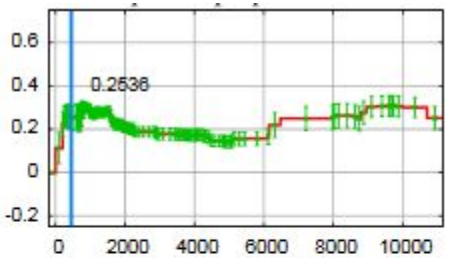


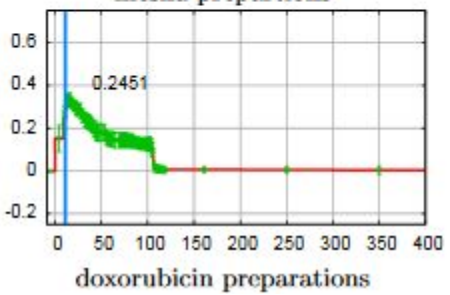


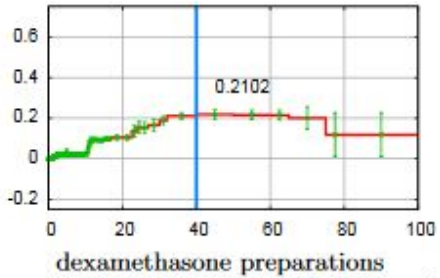
PNEUMONIA RISK

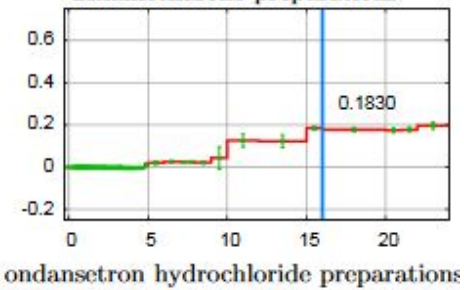
Paciente 2:





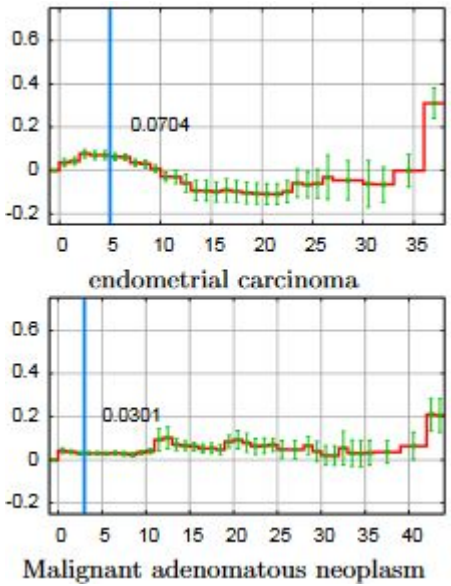


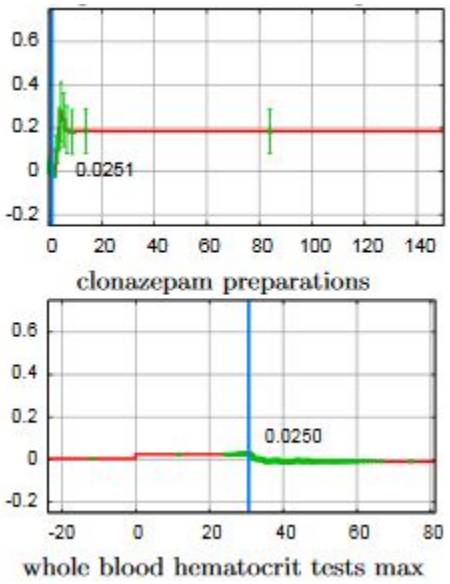


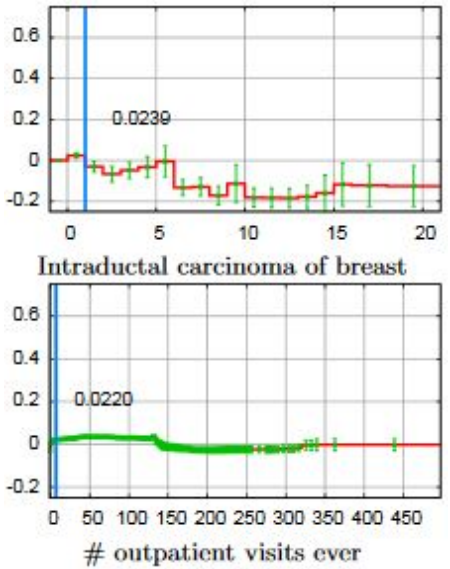


PNEUMONIA RISK

Paciente 3:







MESTRADO
ENGENHARIA ELETRÔNICA
E COMPUTAÇÃO

RODRIGO BLANKE LAMBRECHT

rodrigoucpel@hotmail.com
Pelotas, 2018

UNIVERSIDADE
CATÓLICA
DE PELOTAS