



# TÉCNICAS DE RUÍDO DE CLASSE PARA APRENDIZADO DE MÁQUINA JUSTO

Ygor de Mello Canalli

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientadores: Geraldo Zimbrão da Silva  
Filipe Braida do Carmo

Rio de Janeiro  
Fevereiro de 2011

TÉCNICAS DE RUÍDO DE CLASSE PARA APRENDIZADO DE MÁQUINA  
JUSTO

Ygor de Mello Canalli

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Orientadores: Geraldo Zimbrão da Silva  
Filipe Braida do Carmo

Aprovada por: Prof. Leandro Guimarães Marques Alvim  
Prof. Nome do Segundo Examinador Sobrenome  
Prof. Nome do Terceiro Examinador Sobrenome  
Prof. Nome do Quarto Examinador Sobrenome  
Prof. Nome do Quinto Examinador Sobrenome

RIO DE JANEIRO, RJ – BRASIL  
FEVEREIRO DE 2011

de Mello Canalli, Ygor

Técnicas de Ruído de Classe para Aprendizado de Máquina Justo/Ygor de Mello Canalli. – Rio de Janeiro: UFRJ/COPPE, 2011.

XI, 9 p.: il.; 29,7cm.

Orientadores: Geraldo Zimbrão da Silva

Filipe Braidão do Carmo

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2011.

Referências Bibliográficas: p. 8 – 8.

1. Aprendizado de Máquina Justo. 2. Ruído em Aprendizado de Máquina. 3. Terceira palavra-chave. I. Zimbrão da Silva, Geraldo *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

*A alguém cujo valor é digno  
desta dedicatória.*

# Agradecimentos

Gostaria de agradecer a todos.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

TÉCNICAS DE RUÍDO DE CLASSE PARA APRENDIZADO DE MÁQUINA  
JUSTO

Ygor de Mello Canalli

Fevereiro/2011

Orientadores: Geraldo Zimbrão da Silva

Filipe Braidão do Carmo

Programa: Engenharia de Sistemas e Computação

Apresenta-se, nesta tese, uma abordagem ...

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

THESIS TITLE

Ygor de Mello Canalli

February/2011

Advisors: Geraldo Zimbrão da Silva

Filipe Braidão do Carmo

Department: Systems Engineering and Computer Science

In this work, we present ...

# Sumário

<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xi</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Justiça em aprendizado de máquina</b>	<b>2</b>
2.1 Justiça, transparência, auditabilidade e privacidade . . . . .	2
2.2 Fontes de injustiça . . . . .	2
2.3 Definições e métricas . . . . .	2
2.4 Regulamentação . . . . .	2
2.5 Conjuntos de dados . . . . .	2
2.6 Abordagens para classificação justa . . . . .	2
2.7 Otimização multiobjetivo . . . . .	2
<b>3 Ruído em aprendizado de máquina</b>	<b>3</b>
3.1 Definições e taxonomia de ruído . . . . .	3
3.2 Classificação na presença de ruído de classe . . . . .	3
3.3 Fatoração e correção de funções de custo . . . . .	3
3.4 Estimadores da matriz de transição . . . . .	3
<b>4 Proposta</b>	<b>4</b>
4.1 Ruído e Injustiça . . . . .	4
4.2 Função de custo de transição para aprendizado justo . . . . .	4
4.3 Correlação e preditores indiretos . . . . .	4
4.4 Regularização . . . . .	4
4.5 Ajuste de hiperparâmetros e matrizes de transição . . . . .	4
<b>5 Resultados e discussão</b>	<b>5</b>
5.1 Significância para redes profundas . . . . .	5
5.2 Estudo comparativo da função de custo de transição . . . . .	5
5.3 Estudo comparativo com regularização . . . . .	5



<b>6</b>	<b>Conclusões</b>	<b>6</b>
<b>7</b>	<b>Conclusões</b>	<b>7</b>
	<b>Referências Bibliográficas</b>	<b>8</b>
<b>A</b>	<b>Algumas Demonstrações</b>	<b>9</b>

# Lista de Figuras

# Lista de Tabelas

# Capítulo 1

## Introdução

# Capítulo 2

## Justiça em aprendizado de máquina

- 2.1 Justiça, transparência, auditabilidade e privacidade
- 2.2 Fontes de injustiça
- 2.3 Definições e métricas
- 2.4 Regulamentação
- 2.5 Conjuntos de dados
- 2.6 Abordagens para classificação justa
- 2.7 Otimização multiobjetivo

## Capítulo 3

# Ruído em aprendizado de máquina

3.1 Definições e taxonomia de ruído

3.2 Classificação na presença de ruído de classe

3.3 Fatoração e correção de funções de custo

3.4 Estimadores da matriz de transição

# Capítulo 4

## Proposta

4.1 Ruído e Injustiça

4.2 Função de custo de transição para aprendizado justo

4.3 Correlação e preditores indiretos

4.4 Regularização

4.5 Ajuste de hiperparâmetros e matrizes de transição

# Capítulo 5

## Resultados e discussão

5.1 Significância para redes profundas

5.2 Estudo comparativo da função de custo de transição

5.3 Estudo comparativo com regularização



# Capítulo 6

## Conclusões

# Capítulo 7

## Conclusões

## Referências Bibliográficas

# Apêndice A

## Algumas Demonstrações