

I Workshop
Sinergia LNCC-empresas
A peça que faltava para a
inovação na sua empresa
www.sinergia.lncc.br

DEXL LAB
DATA EXTREME LAB

Fabio Porto, fporto@lncc.br

Projetos em Andamento



Acompanhamento e Predição de Atletas de alto rendimento.
Sistema SAHA: Apoio Holístico ao Atleta

Sistema Gypscie – G e Seleção de Modelo Predição - Petrobras

Plataforma de Big Data para Saúde (PDCaS) – FIOCRUZ ICICT

Modelos de IA para detecção de Nichos Ecológicos;

Análise de Redes Complexas, Redes de Telefonia Celular; Redes Dinâmicas no Tempo.

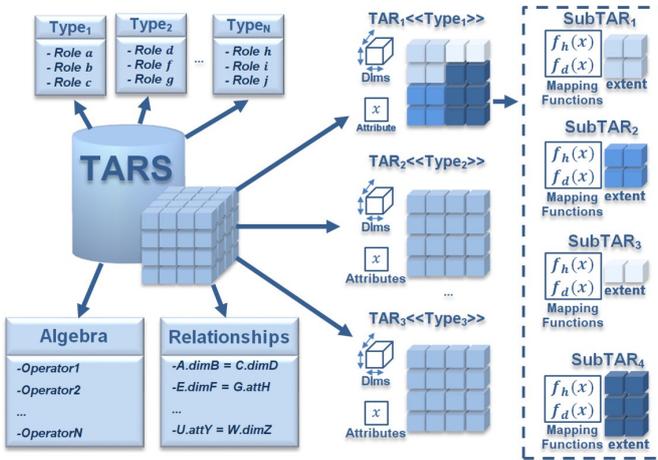
DEXL group

Predição de Meteorologia Usando Redes de Convolução LSTM (USP- IAG)

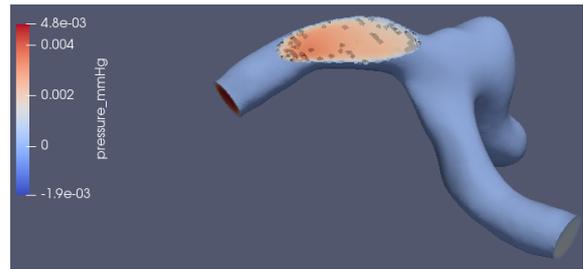
Busca por Lentes Gravitacionais: Apache Spark + DES/SDSS - LineA

DEXL LAB
DATA EXTREME LAB

Sistemas – SAVIME



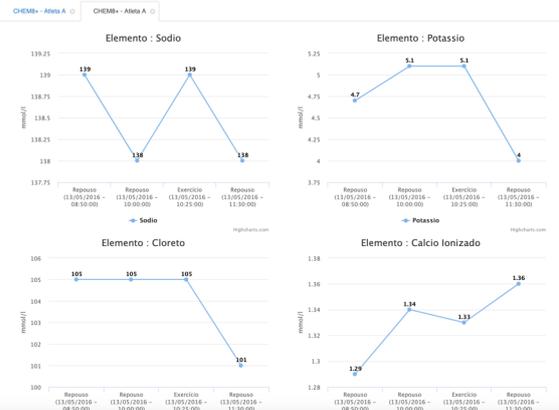
CATALYZE(
SUBSET(DERIVE(HemolabData, pre_mmHg,
Pressure*0.00750062), time, 0, 20),
WHERE(Geometry, coords:2 < -13.0), Topology,
"/home/foo/outputdir/")



DEX LAB
 DATA EXTREME LAB

Tese de Doutorado Hermano L. Lustosa

Sistemas - SAHA



Sistemas: Data Science Suite





TextAnalysis

TextAnalysis
Application which lets the user explore the contents of a PDF file and compare two different PDFs for similar subjects



Data Visualization
Understand your data

DataExploration
Application which lets the user do data exploration on a tabular dataset with a interface



Plant Classification

Plant_Classification
Application to classify plant images into possible species and classify their health status



GraphAnalysis

GraphAnalysis
An Application that receives the adjacency matrix of a graph in a tabular form and lets the user see many metrics of the graph



SentimentAnalysis



Deep Learning Image Predictor



TimeSeries



Machine Learning



OBJECT DETECTION
Find them all

ObjectDetection
An application that receives a image and detects all objects trained on the coco dataset



Sistemas: Plataforma BigData para Saúde Cooperação FIOCRUZ/ICICT






Data Science Lab
Inovação e Aprendizagem Colaborativa





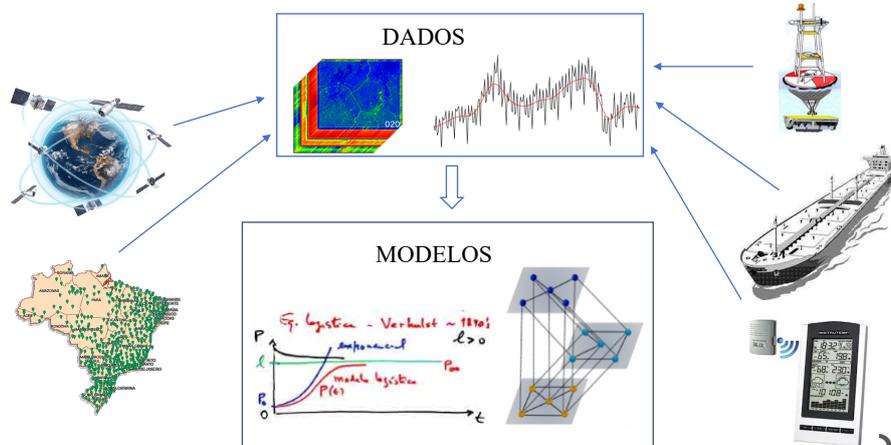
Previsão Espaço-Temporal

O Caso de Uso da Meteorologia – Tese Doutorado Yania Molina Souto

DEXL LAB
DATA EXTREME LAB



- Aumento da disponibilidade de dados espaço – temporais.
- Custos cada vez menores para acessar os dados.
- Modelos baseados em dados cada vez mais precisos.



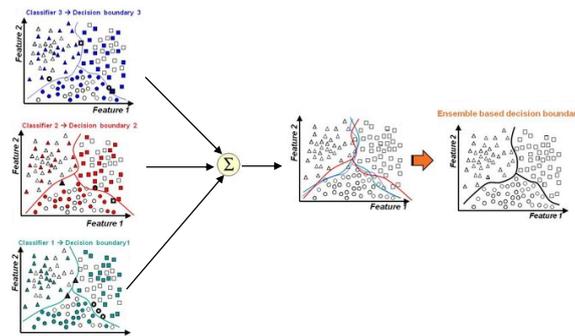
Junho, 2018 | LNCC

DEXL LAB
DATA EXTREME LAB

The image shows two weather forecast interfaces for Rio de Janeiro. The left interface is from CLIMATEMPO, displaying a current temperature of 25°C and a forecast for today with a high of 28°C and a low of 16°C. The right interface is from The Weather Channel, showing a current temperature of 26°C and a forecast for today with a high of 26°C and a low of 18°C. Both interfaces include hourly temperature forecasts for the afternoon and evening.

Métodos de previsão em conjunto

Dado um conjunto X de exemplos de treinamento e um conjunto $H = \{h_1, h_2, \dots, h_n\}$, onde cada h_i é uma hipótese (modelo) de uma função real desconhecida f , e um método de aprendizagem em conjunto E , é possível gerar uma predição $y_i = E(h_i)$, $1 \leq i \leq n$, ponderando as saídas o_i dos n modelos de base.



Fonte: R. Polikar, "Ensemble based systems in decision making," IEEE Circuits and Systems Magazine, vol. 6, no. 3, pp. 21–45, 2006.



Adoção da rede ConvLSTM + Comitê

Junho, 2018 | LNCC

13 / 50



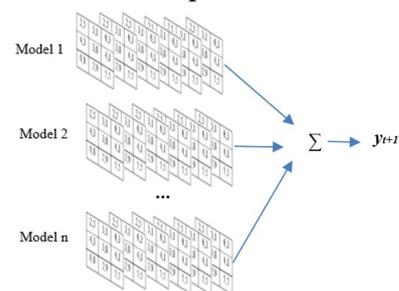
Baseados na natureza dos dados em problemas de aprendizagem em conjunto espaço-temporal, adota-se uma arquitetura ConvLSM.

Vantagens:

- Incorpora a semântica temporal dos dados no aprendizado.
- Incorpora a semântica espacial dos dados no aprendizado.
- Permite guardar as relações espaciais entre cada padrão espacial aprendido.
- É possível definir fronteiras e métodos para tratá-las.

Limitações:

- Análise de imagem e vídeo.
- Não permite análise de mais de uma série simultaneamente.

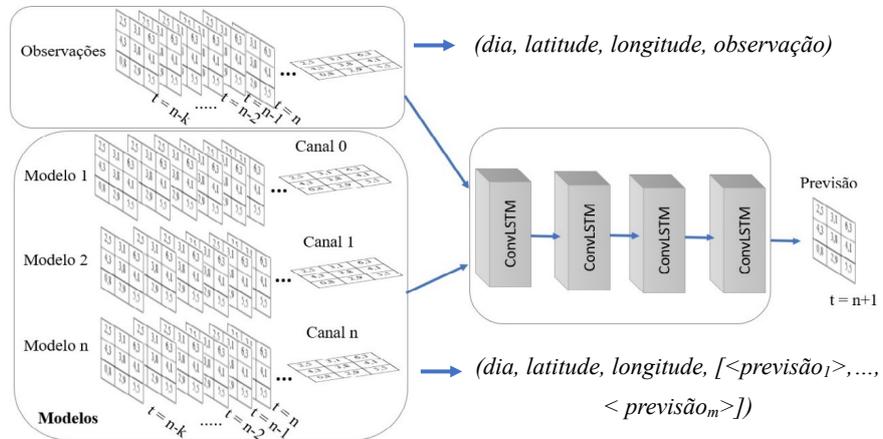


Junho, 2018 | LNCC

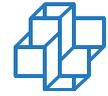
DEXLAB
DATA EXTREME LAB



Usando os canais para representar os modelos



Resultados Experimentais



Caso de Estudo I: Chuva

Junho, 2018 | LNCC



Região Espacial para Predição

(a) Dados de observação



(b) Previsões dos modelos

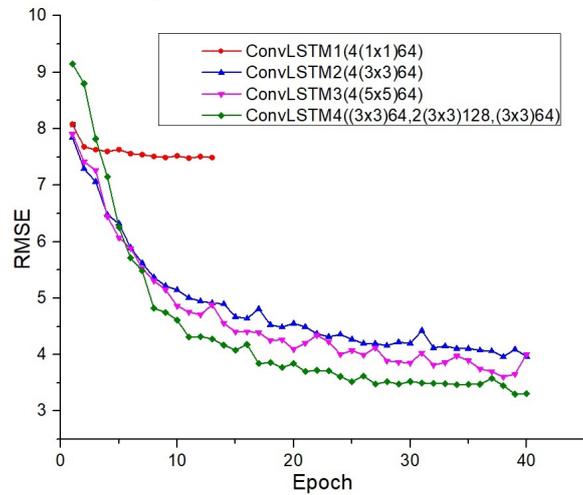


Junho, 2018 | LNCC





Complexidade da arquitetura

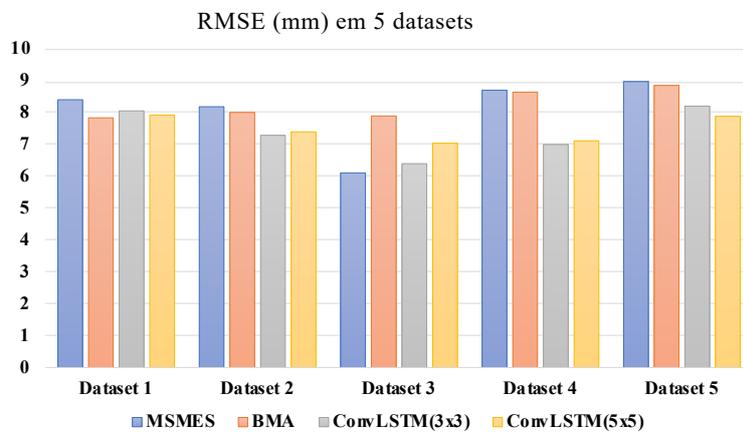


Experimento inicial testando os modelos em 5% das amostras.

Junho, 2018|LNCC



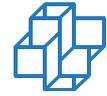
Comparação com outros métodos



RMSE médio de cada método em toda a grade bidimensional.

Junho, 2018|LNCC





Resumo de erros em 5 datasets (70/15/15)

Modelo	RMSE(D1)	RMSE(D2)	RMSE(D3)	RMSE(D4)	RMSE(D5)
BMA	7.85	8.02	7.91	8.66	8.87
MSMES	8.42	8.20	6.12	8.72	8.99
ConvLSTM(3x3)	8.07	7.30	6.41	7.01	8.22
ConvLSTM(5x5)	7.94	7.41	7.05	7.13	7.90

RMSE(mm) médio de cada método em toda a grade bidimensional.

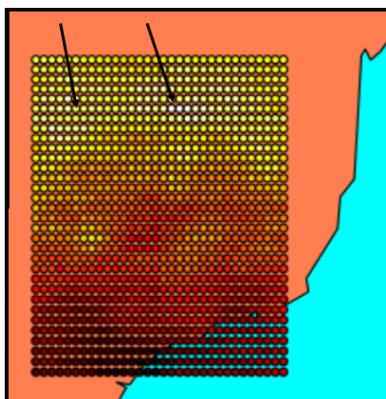
Junho, 2018|LNCC

DEXLAB
DATA EXTREME LAB

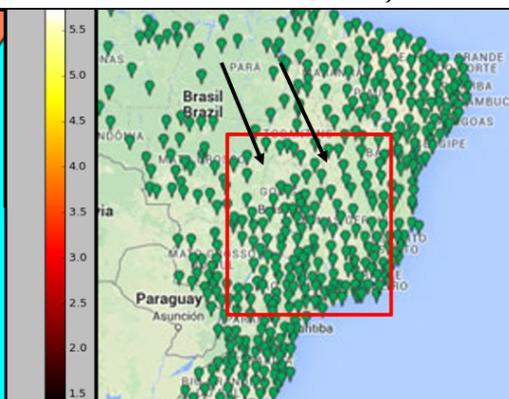


ConvLSTM(4(5x5)64) MAE = 3.86 RMSE = 7.83

Erros maiores



Pouca informação



MAE médio (mm) (24 horas)

Junho, 2018|LNCC

DEXLAB
DATA EXTREME LAB

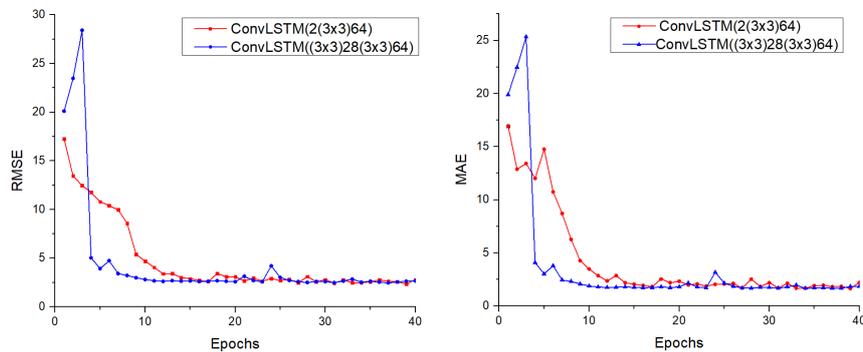


Caso de Estudo II: Temperatura

Junho, 2018|LNCC

DEXLAB
 DATA EXTREME LAB

Complexidade da arquitetura



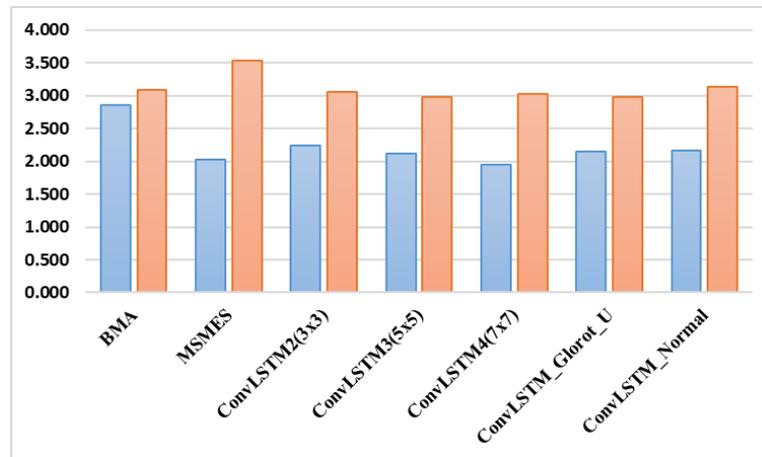
Modelo	MAE(°C)	RMSE(°C)
ConvLSTM(2(3x3)64)	2.247	3.059
ConvLSTM((3x3)28(3x3)64)	2,089	3,140

- Erros médios ao testar os modelos na grade de (32 x 32).
- Modelos gerados em 40 iterações.

Junho, 2018|LNCC

DEXLAB
 DATA EXTREME LAB

Comparação com outros métodos



MAE (azul) e RMSE (vermelho) dos modelos em °C

Junho, 2018|LNCC

DEXLAB
DATA EXTREME LAB

Comparação com outros métodos

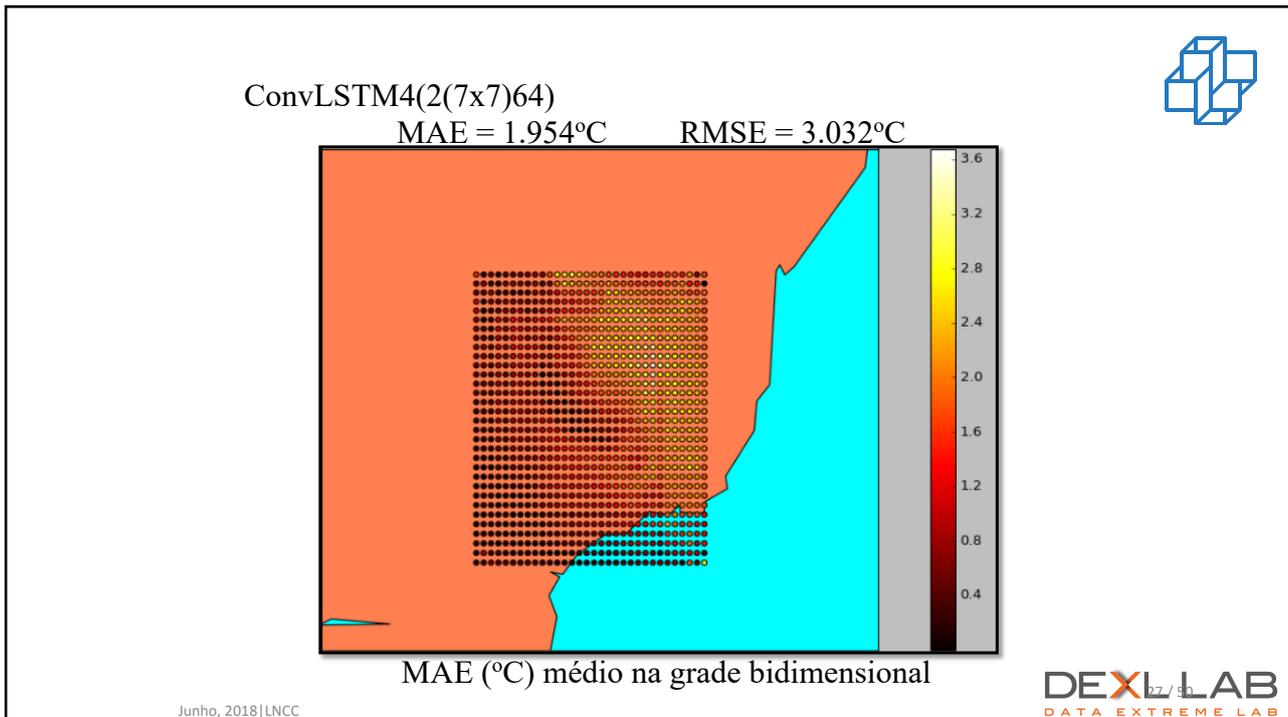


Modelo	MAE(°C)	RMSE(°C)
BMA	2.864	3.092
MSMES	2.023	3.540
ConvLSTM2(3x3)	2.247	3.059
ConvLSTM3(5x5)	2.122	2.976
ConvLSTM4(7x7)	1.954	3.032
ConvLSTM_Glorot_U	2.144	2.986
ConvLSTM_Normal	2.159	3.129

A nova versão da rede, com nova arquitetura e usando um número maior de iterações leva a um erro da ordem de aprox. 0.27.

Junho, 2018|LNCC

DEXLAB
DATA EXTREME LAB



Comentários Finais

- DEXL : Big Data; IA; Data Science
- Softwares Desenvolvidos: SAHA; SAVIME; DSS
- Previsão de Temperatura e Chuva
 - Redes Profunda CONVLSTM – Espaço-Tempo
 - Técnica de comitê de modelos físicos
 - Predição de temperature com erro approx. 0.027
 - Modelo sendo avaliado por equipe da Clima-Tempo

Data EXtreme Lab (DEXL)



Dr. Fabio Porto
fabio.porto@dbs2.com.br





Muito Obrigado

Fabio Porto

fporto@linc.br



PETROBRAS



cidacs
Centro de Integração de Dados
e Conhecimentos para Saúde
FIOCRUZ



Laboratório
Nacional de
Computação
Científica

DEXL LAB
DATA EXTREME LAB