CS109 Project Proposal

Title

\*Obrigatório

Title of your proposal \*

Previsão da Incidência de Dengue em Capitais Brasileiras

Your Group

Enter here the names and emails of all the members of your group. You MUST have at least 3 members, and at most 4 members in your group, unless you are a DCE student or (due to a very good reason) have received permission otherwise from the instructors.

Group Member 1's Name \*

Lucas Ribeiro Coutinho

Group Member 1's email \*

Enter your email. This should be the email you used for your course registration.

lucasribeirocoutinho@gmail.com

Group Member 2's Name

Valter Moreno

Group Member 2's email.

Enter your email. This should be the email you used for your course registration.

valter.moreno@gmail.com

Group Member 3's Name

Group Member 3's email

Enter your email. This should be the email you used for your course registration.

Group Member 4's Name

Group Member 4's email

Enter your email. This should be the email you used for your course registration.

About your Project

Background and Motivation \*

Discuss your motivations and reasons for choosing this project, especially any background or research interests that may have influenced your decision.

Vários fatores nos incentivaram a estudar o tema escolhido: (1) a importância do tema para a realidade brasileira vem crescendo nas últimas décadas, sendo os custos sócio-econômicos envolvidos na prevenção e tratamento da doença substanciais; (2) temos acesso à base de dados do projeto Infodengue (Codeço et al., 2016), que contém dados confiáveis e atualizados sobre a ocorrência da doença no Brasil desde 2008; (3) o volume de dados disponíveis proporciona uma oportunidade para aplicar diversas técnicas de análise e exercitar o uso de tecnologias de computação distribuída.

Project Objectives \*

What are the scientific and inferential goals for this project? What would you like to learn and accomplish? List the benefits.

O objetivo principal do trabalho é avaliar a eficácia de diferentes modelos de previsão da incidência da dengue ao longo do tempo em função da proximidade entre as capitais analisadas. Isso nos dará a oportunidade de nos aprofundarmos em técnicas de machine learning como regressão, random forests, e redes neurais (LSTM), além de técnicas de visualização de dados geoespaciais.

Must-Have Features \*

These are features or calculations without which you would consider your project to be a failure.

(1) visualização dos dados geoespaciais e sua variação ao longo do tempo;

(2) cálculo das correlações cruzadas das séries temporais estacionárias de ocorrência de dengue;

(3) obtenção de modelos baseados nas técnicas acima citadas;

(4) avaliação da capacidade de previsão dos modelos gerados.

Optional Features

Those features or calculations which you consider would be nice to have, but not critical.

What Data? \*

From where and how are you collecting your data?

O conjunto de dados usado no artigo foi fornecido pelo projeto InfoDengue (Codeço et al., 2016). Confome descrito por Mussumeci (2018), o projeto, desenvolvido em parceria entre a Fundação Oswaldo Cruz e a Fundação Getúlio Vargas, implementou um sistema integrado de alerta de dengue, com uma fonte única de dados cuidadosamente selecionados para estudos epidemiológicos. O InfoDengue monitora, até 2018, 790 cidades no Brasil, prevendo o número de novos casos de dengue em cada um deles semanalmente.

Design Overview \*

List the statistical and computational methods you plan to use.

(1) correlações cruzadas de séries temporais;

(2) visualização de dados geoespaciais e temporais;

(2) técnicas de machine learning (regressão, random forest, LSTM)

Verification \*

How will you verify your project's results? In other words, how do you know that your project does well?

(1) Os resultados da análise exploratória e visualização de dados serão avaliados por um dos líderes do projeto Infodengue, Prof. Flavio Codeço.

(2) Seguindo o processo tradicional de desenvolvimento de modelos preditivos de machine learning, os dados serão divididos em amostras de treinamento, validação e teste, usando técnicas de k-fold cross-validation.

(3) A capacidade preditiva dos modelos será avaliada por meio do cálculo de indicadores tais como RMSE.

Visualization & Presentation \*

How will you visualize and communicate your results?

Os resultados vão ser apresentados num website, em linguagem gráfica e textual orientada ao público consumidor das informações sobre a dengue (pesquisadores e sociedade em geral). Eles incluirão gráficos interativos das séries temporais de ocorrência da doença nas capitais brasileiras, e mapas com informação da variação ao longo do tempo da ocorrência da dengue. Além disso, será feita uma comparação da efetividade dos modelos desenvolvidos.

Schedule \*

Make sure that you plan your work so that you can avoid a big rush right before the final project deadline, and delegate different modules and responsibilities among your team members. Write this in terms of weekly deadlines.

As seguintes datas de entrega estão sendo propostas:

09/08 - Análise exploratória e visualização de dados concluídas

23/08 - Modelos desenvolvidos e testados

30/08 - Website publicado

06/09 - Homologação dos resultados gerados pelo representante do Infodengue

13/09 - Entrega da versão final

Project due date

Your project is due on December 12 (no extensions). Make sure your schedule above reflects this!

Nunca envie senhas pelo Formulários Google.

Este conteúdo não foi criado nem aprovado pelo Google. Denunciar abuso - Termos de Serviço

Google Formulários