

Atividade 02 - EBAC

Lucas Amaral Taylor

30 de dezembro de 2024

1 Explique com suas palavras o conceito de Machine Learning.

Machine Learning, em tradução livre, quer dizer “aprendizado de máquina” e trata-se de um método em que se ensina à máquina determinada atividade a partir de uma base de dados na qual se encontram padrões.

2 Diferencie modelos descritivos de modelos preditivos.

Primeiro, para diferenciar os dois modelos é necessário atentar-se aos nomes de cada um. O modelo preditivo visa realizar uma previsão, ou seja, a partir da análise de dados existentes tenta-se estimar um valor de interesse. Enquanto isso, o modelo descritivo tem como objetivo descrever o conjunto de dados existentes através das características que os dados compartilham entre si, em outras palavras, o modelo descritivo analisa padrões.

Além disso, temos que o modelo preditivo é um modelo supervisionado, enquanto o modelo descritivo é **não** supervisionado. Lembrando que um modelo supervisionado é aquele que utiliza como referência dados históricos.

3 Pesquise aplicações de modelos de Machine learning voltados para sua área de interesse e descreva exemplificando o que encontrou.

Ainda não tenho uma área profissional de interesse definida, porém, em breve, farei um curso da área de meteorologia e acabei baseando minha pesquisa nesta área.

Nela, encontrei o artigo intitulado “*Machine Learning Methods in Weather and Climate Applications: A Survey*” [1], o artigo analisa modelos estatísticos¹ de previsão do tempo, compara com os modelos físicos em fenômenos de curto e médio-longo prazo.

Como citado no resumo do artigo, em tradução-livre, temos que:

¹Os modelos estatísticos citados no artigo utilizam-se de *Machine Learning*

“Esta pesquisa tem como objetivo consolidar o entendimento atual das aplicações de Aprendizado de Máquina na previsão do tempo e do clima - um campo de importância crescente em vários setores, incluindo agricultura e gerenciamento de desastres. (...) De acordo com a pesquisa, embora o Aprendizado de Máquina demonstre recursos significativos na previsão do tempo de curto prazo, sua aplicação na previsão do clima de médio a longo prazo permanece limitada, restringida por fatores como variáveis climáticas complexas e limitações de dados.”[1]

4 Escolha um das aplicações que citou acima e descreva como você imaginaria que seriam aplicadas as etapas do Crisp-DM a esse projeto.

No artigo, são aprofundados oito modelos meteorológicos. A fim de restringir-me à atividade, relacionarei apenas um com o Crisp-DM. O modelo selecionado foi o **ClimaX**² e relacionaremos com as etapas do Crisp-Dm apresentadas em aula:

1. Entendimento e Definição de Objetivos.

Na meteorologia, geralmente usa-se métodos numéricos que baseiam-se em leis físicas. Tais métodos, tem como características uma dinâmica não linear, interações complexas em múltiplas variáveis e computacionalmente exaustivos. O ClimaX apresenta-se como um modelo de aprendizado profundo flexível e generalizado para a meteorologia, fundamentando a partir de dados meteorológicos e com menos gasto computacional.

Em suma, o ClimaX tem como objetivo ser uma alternativa mais eficiente para a análise de fenômenos do clima e tempo.

2. Preparação dos Dados.

Os dados do modelo são originários do *Coupled Model Intercomparison Project*, CMIP. Em tradução livre, temos que:

O CMIP é um projeto do Programa Mundial de Pesquisa Climática (WCRP) que fornece projeções climáticas para entender as mudanças climáticas passadas, presentes e futuras. O CMIP e sua infraestrutura de dados associada tornaram-se essenciais para o Painel Intergovernamental sobre Mudanças Climáticas (IPCC) e outras avaliações climáticas internacionais e nacionais. [5]

No que diz respeito a exploração, os autores de [4] comentam que: “No entanto, lidar com o conjunto inconsistente de variáveis em diferentes fontes de dados pode ser um desafio” e realizam um filtro na fase de exploração, detalhado devidamente na seção 3 de [4]. Por fim, no que diz respeito ao tratamento dos dados, os autores optaram por reanalisar os dados a partir de outra base de dados, ERA5 [2].

3. Modelagem.

Na seção 4 de [4], os autores levantam os seguintes pontos:

Nós realizamos o *fine-tuning*³ do ClimaX em um conjunto diversificado de tarefas para avaliar seu desempenho e generalidade.

²As informações foram retiradas de “*ClimaX: A foundation model for weather and climate*” [4].

³O *fine-tuning* é o processo de ajustar um modelo treinado para melhorar seu desempenho em uma tarefa específica.

As tarefas são categorizadas em *previsão*, *projeção climática* e *downscaling climático*⁴. Os experimentos têm como objetivo responder às seguintes questões:

- (a) Como o ClimaX se comporta em previsão global em comparação com o atual sistema de NWP⁵?
- (b) Podemos realizar o *fine-tuning* do ClimaX para fazer previsões para uma região específica ou em diferentes horizontes temporais a partir do pré-treinamento?
- (c) Quão bem o ClimaX se sai em tarefas climáticas que são completamente diferentes do pré-treinamento?

Além dos experimentos principais, analisamos o desempenho do ClimaX melhora com o aumento do tamanho dos dados, da capacidade do modelo e da resolução dos dados.

4. Avaliação.

Finalmente, a avaliação do modelo consiste em comparações e análises constantes com modelos já consolidados na área, vide [4] e [1] e, claro, com a realidade.

Nota-se que, segundo [1], o ClimaX destaca-se em previsão de temperatura e eventos extremos em um curto período de tempo. Além disso, por se tratar de um modelo novo, 2023, há muitos testes e refinamento do modelo a serem realizados pela comunidade científica.

A Disponibilidade do modelo

O modelo possui um repositório disponível no *Github*, disponível aqui.

⁴O *downscaling climático* é a técnica de melhorar previsões climáticas de grandes escalas para escalas regionais ou locais.

⁵“Os dados da Previsão Numérica do Tempo (NWP) são a forma mais conhecida de dados de modelos meteorológicos.” [3]

Referências

- [1] Liuyi Chen, Bocheng Han, Xuesong Wang, Jiazhen Zhao, Wenke Yang, and Zhengyi Yang. Machine learning methods in weather and climate applications: A survey. *Applied Sciences*, 13(21), 2023.
- [2] Copernicus Climate Change Service (C3S). Reanalysis era5: Single levels, 2024. Acesso em: 29 dezembro 2024.
- [3] National Centers for Environmental Information (NCEI). Numerical weather prediction (nwp), 2024. Acesso em: 29 dezembro 2024.
- [4] Tung Nguyen, Johannes Brandstetter, Ashish Kapoor, Jayesh K. Gupta, and Aditya Grover. Climax: A foundation model for weather and climate, 2023.
- [5] World Climate Research Programme. Cmp: Coupled model intercomparison project, 2024. Accessed: 2024-12-29.