



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

ADRIANO MARABUCO DE ALBUQUERQUE LIMA

Sistema baseado em seleção dinâmica para previsão de casos de COVID-19

Recife

2022

ADRIANO MARABUCO DE ALBUQUERQUE LIMA

Sistema baseado em seleção dinâmica para previsão de casos de COVID-19

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação.

**Área de Concentração:** Inteligência Computacional

**Orientador (a):** Paulo S. G. de Mattos Neto

Recife

2022

Catálogo na fonte  
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

L732s Lima, Adriano Marabuco de Albuquerque  
Sistema baseado em seleção dinâmica para previsão de casos de COVID-19 / Adriano Marabuco de Albuquerque Lima. – 2022.  
59 f.: il., fig., tab.

Orientador: Paulo S. G. de Mattos Neto.  
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2022.  
Inclui referências.

1. Inteligência computacional. 2. Séries temporais. 3. Aprendizado de máquina. I. Mattos Neto, Paulo S. G. de (orientador). II. Título.

006.31

CDD (23. ed.)

UFPE - CCEN 2023-28

**Adriano Marabuco de Albuquerque Lima**

**“Sistema baseado em seleção dinâmica para previsão de casos de COVID-19”**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação. Área de Concentração: Inteligência Computacional.

Aprovado em: 13/12/2022.

**BANCA EXAMINADORA**

---

Prof. Dr. Adriano Lorena Inacio de Oliveira  
Centro de Informática / UFPE

---

Prof. Dr. João Fausto Lorenzato de Oliveira  
Escola Politécnica / UPE

---

Prof. Dr. Paulo Salgado Gomes de Mattos Neto  
Centro de Informática / UFPE  
**(Orientador)**

Dedico esta dissertação à minha família, amigos e mestres.

## AGRADECIMENTOS

Agradeço à Providência Divina pela oportunidade de superar este desafio enfrentando probabilidades adversas, e aprendendo com os erros no vasto espaço de busca da Computação.

Este trabalho não seria possível sem a colaboração de algumas pessoas em momentos críticos da série temporal que resultou nesta dissertação.

Agradeço a minha família e, em especial a minha mãe, pela valorização da educação, pelo exemplo e apoio incondicional aos meus projetos pessoais, sempre acreditando no meu potencial.

Agradeço a minha esposa pelo companheirismo e incentivo, que me ajudaram a sonhar com o Mestrado e atravessar as percalços desta jornada.

Agradeço ao Prof. Paulo Salgado Gomes de Mattos Neto pelo entusiasmo e liberdade na pesquisa, assim como, pela crítica construtiva e direcionamento para manutenção do foco e qualidade nos trabalhos.

Agradeço ao colega Jair Paulino de Sales pela parceria durante a realização das disciplinas e pesquisas.

Agradeço à comunidade do Centro de Informática pela acolhida, ensinamentos e pertencimento a uma instituição de ensino e pesquisa de excelência.

*"The most that can be expected from any model is that it can supply a useful approximation to reality: All models are wrong; some models are useful." (BOX et al., 1978, p. 440).*

## RESUMO

A pandemia da COVID-19 provocou 546 milhões casos e 20 milhões óbitos até 30 de junho de 2022, além disso gerou uma queda de 3% no PIB mundial em 2020. Um dos desafios no enfrentamento da doença é a previsão da quantidade de casos e óbitos, assim como a tendência de crescimento e decrescimento. Muitos trabalhos têm focado nessa tarefa, entretanto nenhum deles considerou a mudança de conceito nas séries temporais da COVID-19. Essas séries possuem basicamente três conceitos principais: crescimento exponencial, decrescimento e platô. Dessa forma, modelos que não tratam dessa questão podem apresentar baixa precisão em virtude da mudança na distribuição dos dados ao longo do tempo. Esse trabalho propõe o *Concept Drift Dynamic Forecasting System (CODYS)* para previsão de casos da COVID-19. O CODYS primeiramente realiza a detecção de conceitos na fase de treinamento para gerar um conjunto de preditores especialistas nos padrões encontrados. Na fase de teste, o método proposto seleciona dinamicamente a partir de uma região de competência o modelo mais apto para prever um dado padrão de teste. O CODYS foi avaliado utilizando conjuntos de dados dez países de diferentes continentes e graus de desenvolvimento. Os resultados mostram que o CODYS alcançou desempenho superior quando comparado com modelos de estatísticos, de aprendizado de máquina e *ensembles*.

**Palavras-chaves:** séries temporais; aprendizado de máquina; mudança de conceito; seleção dinâmica; COVID-19.

## ABSTRACT

The COVID-19 pandemic caused 546 million cases and 20 million deaths by June 30, 2022, in addition to generating a 3% drop in world GDP in 2020. One of the challenges in coping with the disease is predicting the number of cases and deaths, as well as the trend of growth and decrease. Many works have focused on this task, however none of them considered the concept change in COVID-19 time series. These series basically have three main concepts: exponential growth, decrease and plateau. Thus, models that do not address this issue may have low accuracy due to the change in data distribution over time. This work proposes the *Concept Drift Dynamic Forecasting System* (CODYS) for forecasting COVID-19 cases. CODYS first performs concept detection in the training phase to generate a set of expert predictors based on the patterns found. In the test phase, the proposed method dynamically selects from a region of competence the model best able to predict a given test pattern. CODYS was evaluated using datasets from ten countries from different continents and levels of development. The results show that CODYS achieved superior performance when compared to statistical, machine learning and *ensembles* models.

**Keywords:** time series; machine learning; concept drift; dynamic selection; COVID-19.

## LISTA DE FIGURAS

Figura 1 – Casos e óbitos da COVID-19 no Mundo. . . . .	14
Figura 2 – Série temporal de CO2 . . . . .	22
Figura 3 – Tendência da Série temporal de CO2 . . . . .	23
Figura 4 – Sazonalidade da Série temporal de CO2 . . . . .	23
Figura 5 – Resíduo da Série temporal de CO2 . . . . .	24
Figura 6 – Correlograma . . . . .	25
Figura 7 – Estacionariedade . . . . .	25
Figura 8 – Diferenciações . . . . .	26
Figura 9 – Padrões da tendência da COVID-19 no Brasil . . . . .	27
Figura 10 – Tipos de Mudança de Conceito. . . . .	28
Figura 11 – Etapa de treinamento. . . . .	33
Figura 12 – Etapa de teste. . . . .	34
Figura 13 – Separação dos Dados. . . . .	36
Figura 14 – Detecção dos Conceitos dos Países. . . . .	43
Figura 15 – Previsão de CODYS e melhor modelo concorrente. . . . .	47
Figura 16 – Teste de Wilcoxon. . . . .	49
Figura 17 – Boxplot do ranking dos modelos. . . . .	49
Figura 18 – Comparativo entre modelo estático e dinâmico . . . . .	50
Figura 19 – Percentual de seleção de cada modelo do pool. . . . .	51
Figura 20 – Tempo computacional para treinamento e teste dos modelos . . . . .	52

## LISTA DE TABELAS

Tabela 1 – Trabalhos sobre Previsão da COVID-19. . . . .	21
Tabela 2 – Estatísticas dos conjuntos de dados . . . . .	35
Tabela 3 – Relação entre tarefa da previsão por tipo de modelo e conjunto de dados usados. . . . .	36
Tabela 4 – Hiperparâmetros dos Modelos. . . . .	38
Tabela 5 – Resultados do teste de Mann–Whitney U comparando as distribuições de duas janelas adjacentes. . . . .	45
Tabela 6 – Comparação em termos de MAPE do CODYS com modelos concorrentes.	45
Tabela 7 – Diferença percentual entre os modelos. . . . .	46

## LISTA DE SÍMBOLOS

$C_N$	Conjunto de dados
$M_N$	Modelos
$T$	Quantidade de <i>lags</i>
$Z_q$	Valor real do padrão de teste
$\hat{Z}_q$	Valor previsto do padrão de teste
$\hat{y}$	Valor Previsto
$y$	Valor Real

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
1.1	MOTIVAÇÃO	15
1.2	OBJETIVOS	17
1.3	ESTRUTURA	17
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	<b>18</b>
<b>3</b>	<b>REFERENCIAL TEÓRICO</b>	<b>22</b>
3.1	SÉRIES TEMPORAIS	22
<b>3.1.1</b>	<b>Modelos Preditivos</b>	<b>26</b>
3.2	MUDANÇA DE CONCEITO	27
3.3	SELEÇÃO DINÂMICA	30
<b>4</b>	<b>SISTEMA PROPOSTO</b>	<b>32</b>
4.1	ETAPA DE TREINAMENTO	32
4.2	ETAPA DE TESTE	33
<b>5</b>	<b>METODOLOGIA DOS EXPERIMENTOS</b>	<b>35</b>
5.1	DESCRIÇÃO DOS DADOS	35
5.2	MODELOS DE PREVISÃO	37
5.3	PRÉ-PROCESSAMENTO, DETECÇÃO DE CONCEITO E SELEÇÃO DINÂMICA	39
5.4	MÉTRICAS DE DESEMPENHO	40
5.5	ESPECIFICAÇÃO DE SOFTWARE E HARDWARE	41
<b>6</b>	<b>RESULTADOS DOS EXPERIMENTOS</b>	<b>42</b>
6.1	ANÁLISE DOS CONCEITOS DETECTADOS	42
6.2	COMPARATIVO DE DESEMPENHO ENTRE MODELOS	45
6.3	COMPARATIVO ENTRE O MODELO ESTÁTICO E DINÂMICO	50
6.4	DESEMPENHO DA SELEÇÃO DINÂMICA EM RELAÇÃO AO <i>ORACLE</i>	51
6.5	ANÁLISE DO TEMPO COMPUTACIONAL	51
<b>7</b>	<b>CONCLUSÃO</b>	<b>53</b>
	<b>REFERÊNCIAS</b>	<b>55</b>

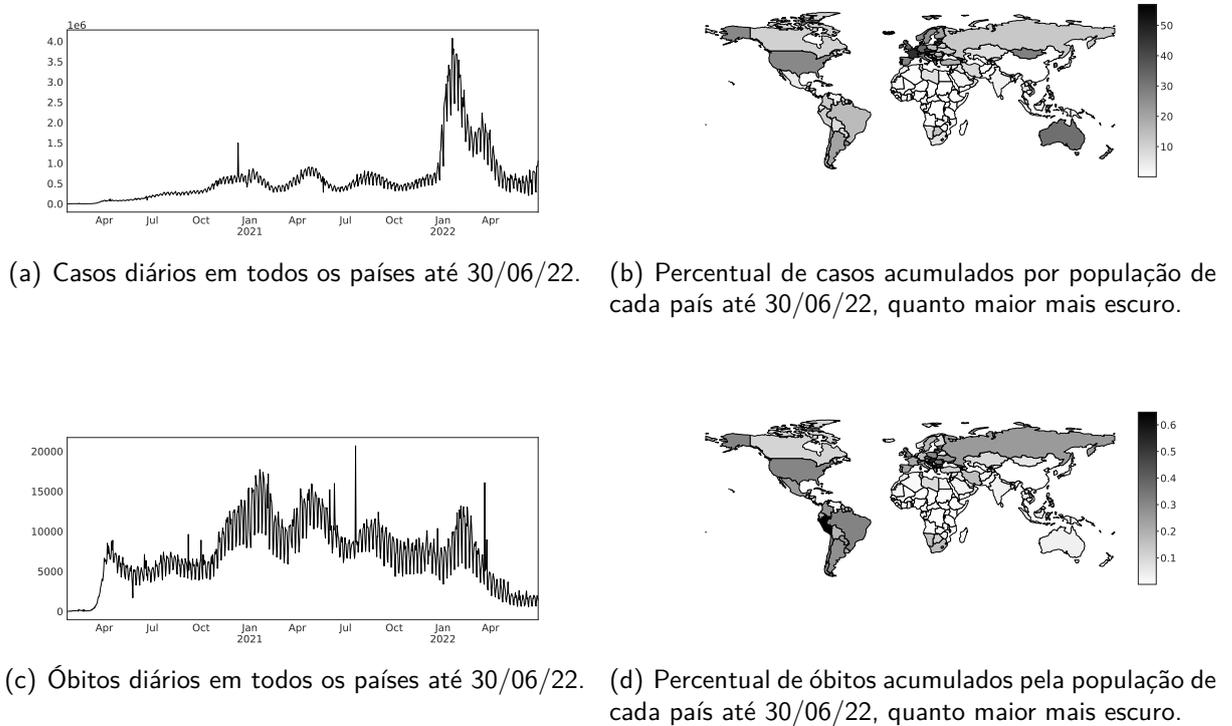
## 1 INTRODUÇÃO

Em dezembro de 2019, uma série de casos severos de pneumonia de razão desconhecida foram reportados em Wuhan, capital da Província de Hubei na China relacionados as vítimas que haviam frequentado o mercado de Frutos do Mar de Huanan. Por volta de 02 de Janeiro de 2020, 41 pacientes tiveram testes positivos para uma nova espécie da família coronavírus (HUANG et al., 2020). Evidências indicam que o vírus originalmente presente em animais, morcegos e pangolins, sofreu mutações e foi transmitido a humanos (HOLMES et al., 2021).

A transmissão do novo coronavírus atingiu rapidamente uma escala global, e em 11 de Março de 2020 pela Organização Mundial de Saúde (OMS) declarou estado de emergência para uma pandemia (CUCINOTTA; VANELLI, 2020). A crise provocada pelo vírus SARS-CoV-2 revelou-se um desafio inédito para as sociedades, governos e sistemas de saúde. Até 30 de junho de 2022 foram contabilizados 546 milhões de casos e 20 milhões de vidas perdidas (DONG et al., 2022).

A evolução dos casos e óbitos apresentou a ocorrência de vários ciclos com momentos de crescimento acelerado na quantidade diária de casos e óbitos, acompanhados por estabilização, que poderia ser seguida de declínio, ou novo crescimento. Além disso, destaca-se a sua abrangência planetária afetando países em todos os continentes e graus de desenvolvimento econômico, conforme se exhibe na Figura 1.

Figura 1 – Casos e óbitos da COVID-19 no Mundo.



**Fonte:** Elaborada pelo autor (2022).

Durante o primeiro ano da pandemia, enquanto não havia vacinas, uma estratégia de enfrentamento da pandemia foi o "achatamento da curva" de casos para evitar o colapso dos sistemas de saúde. Surtos de contágio do vírus e casos graves geram um excesso de demanda por leitos em volume muito superior à capacidade de atendimento. A redução do contágio por meio de políticas públicas restritivas permitiu diminuir o contato entre as pessoas, causando uma menor velocidade de transmissão da doença (SAEZ et al., 2020).

Deve-se considerar, entretanto, que a adoção dessas medidas teve grande impacto na economia. No ano de 2020, a atividade econômica contraiu em 90% dos países, o que resultou em uma queda de 3% no PIB mundial, e aumento na pobreza global pela primeira vez na geração (BANK, 2022). Em termos comparativos, essa perda afetou mais países do que outras grandes crises globais como a grande depressão de 1930, as duas guerras mundiais, a crise dos países emergentes de 1980 e a crise do *subprime* de 2007 a 2009.

Diante da necessidade solucionar os problemas causados pela pandemia, a capacidade de analisar e prever o comportamento da pandemia assume importância fundamental para a tomada de decisão, auxiliando, por exemplo no equilíbrio entre riscos de saúde e econômicos.

## 1.1 MOTIVAÇÃO

A sequência de casos diários da COVID-19 representa uma série temporal, que é um conjunto de dados obtidos sequencialmente no tempo. Uma característica intrínseca das séries temporais é que as observações adjacentes podem apresentar relações de dependência significativas entre si (BOX et al., 2015). Frequentemente as séries temporais são divididas em três componentes: i) a tendência, que representa alterações de valor contínuas e não cíclicas; ii) a sazonalidade, que representa alterações periódicas e cíclicas; iii) resíduo, que representa alterações aleatórias (COWPERTWAIT; METCALFE, 2009), exibidos nas Figuras 3,4 e 5. Desta forma, a previsão do comportamento da pandemia requer a utilização de algoritmos preditivos capazes de modelar componentes e relações de dependência nas séries temporais para poder inferir a quantidade futura de casos, a partir dos dados passados (COMITO; PIZZUTI, 2022).

A previsão da COVID-19 requer algoritmos capazes de lidar com relações de dependência temporal, que mudam com o tempo, por conta da sua complexidade e multifatorialidade (UCHIDA; YOSHIDA, 2022). Dentre os diversos fatores que influenciam as séries pandêmicas, destacam-se as i) medidas restritivas, ii) comportamento social, iii) mutações do vírus e iv) vacinas (HSIANG et al., 2020; MEHTA et al., 2021; REN et al., 2022; BERNAL et al., 2021).

As medidas restritivas como *lockdowns*, fechamento de fronteiras e isolamento social foram impostas para reduzir o contágio da doença. Tais medidas teriam prevenido ou atrasado a infecção de 61 milhões de casos em seis países estudados por (HSIANG et al., 2020), entretanto há estudos que questionam a eficácia dessas medidas (ALLEN, 2022; HERBY; JONUNG; HANKE, 2022).

O comportamento social como o uso de máscaras, higienização das mãos, aglomeração social e eventos influenciavam nos números de casos. Como exemplo, o feriado de ação de graças nos Estados Unidos provocou um aumento na quantidade de viagens e encontros, que levou a um aumento no volume de casos (MEHTA et al., 2021).

Além disso, as mutações do vírus, com taxas de transmissão e letalidade diferentes, resultaram em surtos de crescimento de casos e óbitos. Por exemplo, na Figura 1, é possível perceber um crescimento exponencial na quantidade de casos na transição de 2021 para 2022, correlacionado com a emergência da variante ômicron, que possuía uma taxa de transmissão mais alta do que as mutações predecessoras (REN et al., 2022).

A evolução da vacinação da população e o tipo de imunizante aplicado afetaram decisivamente na quantidade de óbitos pela COVID-19 (BERNAL et al., 2021). Na Figura 1 percebe-se

que a tendência de crescimento dos óbitos ano 2020 sofre uma reversão, após as vacinações em massa, que foram iniciadas no ano 2021, e, em 2022, há uma queda abrupta no volume de óbitos, que atinge o menor patamar desde o início da pandemia.

Os diferentes vetores mencionados, dentre outros, influenciam e alteram a distribuição dos dados da COVID-19 ao longo do tempo gerando mudanças de conceito e tendência. A mudança de conceito (*concept drift*) ocorre quando há uma alteração na distribuição conjunta entre a variável preditora e a variável resposta (SCHLIMMER; GRANGER, 1986; WIDMER; KUBAT, 1996).

Na pandemia é possível identificar três tipos principais de conceitos: crescimento exponencial, platô e declínio, como se observa na Figura 9. No crescimento exponencial há uma tendência positiva com aumento acelerado de casos ou óbitos. No platô não há de tendência. No declínio há uma tendência negativa com redução progressiva na variação de infectados e vítimas. As transições entre conceitos tornam esta tarefa de previsão um problema computacionalmente desafiador.

Ao longo da crise, a comunidade científica debruçou-se sobre o problema e explorou abordagens usando modelos epidemiológicos (SIS, SIR e SEIR), clássicos de séries temporais (ARIMA e ETS), aprendizagem de máquina clássica (SVR, MLP e XGboost) e aprendizado profundo (LSTM e GRU) para lidar com os desafios da previsão (COMITO; PIZZUTI, 2022). Recorreu-se também a utilização de múltiplos preditores e modelos híbridos. No entanto, as soluções propostas não consideraram a mudança de conceito na série temporal, e trataram do problema utilizando um único modelo preditivo ou uma seleção estática de modelos, cujas previsões não eram influenciadas pela dinâmica das mudanças no conjunto de teste. Nesse cenário, identifica-se uma ausência de trabalhos sobre previsão da COVID-19 que lidam com a questão da mudança de conceito, o que representa uma lacuna, a qual se busca preencher com a presente pesquisa.

Adotando como ponto de partida o problema de *concept drift* nas séries pandêmicas, a hipótese de pesquisa é usar uma estratégia de "dividir para conquistar"<sup>1</sup> e treinar diversos preditores especialistas nos diferentes conceitos da pandemia para compor um sistema de múltiplos de preditores. Na etapa de teste para cada padrão de teste é apresentada anteriormente uma região de competência, o sistema prevê os valores para essa região e, com base no menor

<sup>1</sup> A estratégia de "dividir para conquistar" representa uma classe de algoritmos que decompõem recursivamente um problema complexo em partes menores, resolvem esses subproblemas e depois integram as soluções parciais até a solução total (DASGUPTA; PAPADIMITRIOU; VAZIRANI, 2008).

erro, seleciona os melhores K modelos. Assim, o *Concept Drift Dynamic Forecasting System* (CODYS) é um sistema de múltiplos preditores baseado em seleção dinâmica capaz de identificar e se adaptar às mudanças de conceito nas diversas fases dos ciclos da pandemia em diferentes países.

## 1.2 OBJETIVOS

O objetivo geral deste trabalho é contribuir com a pesquisa em séries temporais de COVID-19 apresentando um modelo preditivo que aborda de forma explícita a mudança de conceito combinada com seleção dinâmica para prever séries não estacionárias como aquelas dos casos da COVID-19. Os objetivos específicos são:

1. Definir uma metodologia para detectar conceitos em séries temporais de COVID-19;
2. Utilizar esses conceitos para criar um sistema de múltiplos preditores;
3. Identificar o conceito vigente no conjunto de teste;
4. Selecionar dinamicamente o(s) preditor(es) mais adequado(s) para previsão considerando o conceito vigente.
5. Analisar os resultados obtidos e comparar com preditores da literatura.

## 1.3 ESTRUTURA

Este trabalho está organizado da seguinte forma: o Capítulo 2 apresenta uma visão geral da literatura sobre previsão da COVID-19, o Capítulo 3 define mudança de conceito e seleção dinâmica, o Capítulo 4 explica o funcionamento do sistema proposto, o Capítulo 5 descreve o protocolo experimental, o Capítulo 6 apresenta os resultados e o estudo de caso do Brasil e o Capítulo 7 traz a conclusão.

## 2 TRABALHOS RELACIONADOS

A previsão de casos e óbitos da COVID-19 tem sido objeto de diversos trabalhos, conforme se observa em revisões da literatura (DEVARAJ et al., 2021; COMITO; PIZZUTI, 2022). Os trabalhos podem ser agrupados em diferentes abordagens: modelos epidemiológicos, clássicos de séries temporais, aprendizagem de máquina clássica e aprendizado profundo.

Alguns autores (PENG et al., 2020; NADIM; GHOSH; CHATTOPADHYAY, 2021) abordaram o problema usando modelos epidemiológicos clássicos como *Susceptible-Infectious-Removed* (SIR), *Susceptible-Exposed-Infectious-Removed* (SEIR), *Susceptible-Infectious-Susceptible* (SIS). Propostos inicialmente por (KERMACK; MCKENDRICK, 1927), tais algoritmos buscam reproduzir a dinâmica da séries de infecção, mortalidade e recuperação da doença por meio do relacionamento entre essas séries e da estimativa das suas distribuições estatísticas. Eles são muito populares na literatura de epidemiologia. Esses modelos foram desenvolvidos especificamente para a previsão de epidemias, e portanto incorporam características peculiares deste tipo problema, como o comportamento das taxas de transmissão e mortalidade. A modelagem a partir de conhecimentos *a priori* da epidemiologia é uma vantagem desta abordagem. Entretanto, o ajuste desses modelos a dados reais apresenta grandes desafios, pois muitas vezes as epidemias não atendem às premissas destes modelos, em especial no que se refere às populações, que raramente são homogêneas espacialmente, e as taxas de transmissão variam de acordo com a idade e fatores individuais (FIRMINO et al., 2020).

Determinados trabalhos lidaram com o problema usando variações do *Autoregressive Integrated Moving Average* (ARIMA e SARIMA) (ARUNKUMAR et al., 2021; HERNANDEZ-MATAMOROS et al., 2020) e do *Holt-Winters* (ETS) (CHAURASIA; PAL, 2020), que são modelos estatísticos clássicos. Tais algoritmos buscam modelar as relações temporais entre os dados considerando os instantes de tempo com correlação estatisticamente significativa, além de lidar com tendência e sazonalidade das séries temporais. Diferente dos modelos epidemiológicos, os modelos de séries temporais são agnósticos e possuem muitos domínios de aplicação. Portanto, se adaptam a diferentes cenários e aplicações. Os modelos são parcimoniosos, ou seja, usam poucos parâmetros, o que mitiga o risco de sobreajuste e prescinde da necessidade de uma otimização complexa de hiperparâmetros. Uma limitação destes modelos é que eles não são capazes de modelar relações não lineares entre os dados (OLIVEIRA; SILVA; NETO, 2021).

Outros trabalhos usaram modelos de aprendizado de máquina, como *Support Vector Ma-*

---

*chines* (SVR) (RIBEIRO et al., 2020), *Multilayer Perceptron* (MLP) (ARDABILI et al., 2020), *Extreme Gradient Boosting* (XGboost) (FANG et al., 2022), *Extreme Learning Machine* (ELM) (TALKHI et al., 2021). Tais algoritmos buscam aprender a função preditora pela minimização dos erros de previsão a partir dos dados das variáveis predictoras e variável de resposta. Os modelos de *Machine Learning* são capazes de aprender relações não lineares entre os dados, o que representa uma vantagem em relação aos modelos clássicos de séries temporais (OLIVEIRA; SILVA; NETO, 2021). Uma precaução a ser tomada com estes modelos é quanto ao risco de sobreajuste, pois a flexibilidade e maior quantidade de parâmetros podem induzir que o treinamento memorize os dados, ao invés de aprender, e isso gera uma perda significativa de desempenho na previsão dos dados de teste.

Os recentes avanços em aprendizado profundo também estimularam pesquisas com estes modelos, como Long Short Term Memory (LSTM) (DEVARAJ et al., 2021) e Gated Recurrent Unit (GRU) (ZEROUAL et al., 2020). Tais modelos usam redes neurais com várias camadas ocultas e células de memória para aprender padrões em dados sequenciais. A flexibilidade e robustez desses modelos permite trabalhar com tarefas complexas que envolvem relações não lineares e de longo prazo entre os dados (DEVARAJ et al., 2021). Uma limitação destes modelos é a necessidade de muitos dados para treinamento, sem os quais eles não conseguem aprender os padrões (BOULMAIZ; GUERMOUI; BOUTAGHANE, 2020), e como envolvem uma maior quantidade de hiperparâmetros a especificação dos modelos torna-se mais desafiadora. Tais condições podem levar ao subajuste ou sobreajuste dos modelos, gerando desempenhos insatisfatórios no conjunto de teste.

Por fim, há também pesquisas utilizando modelos híbridos (ALA'RAJ; MAJDALAWIEH; NIZAMUDDIN, 2021; MORAIS; GOMES, 2022) e combinações (*ensembles*). As combinações consistem em modelos independentes cujas previsões são combinadas para obtenção da previsão final (OPITZ; MACLIN, 1999). Um exemplo destes trabalhos foi utilização de um *ensemble* de ARIMA com *Stacked Long Short Term Memory* (S-LSTM) com pesos ponderados (MAALIW et al., 2021). Os modelos híbridos são construídos pela integração entre diferentes sistemas inteligentes (MEDSKER, 2012). Um destes trabalhos utilizou o modelo *Susceptible-Exposed-Infectious-Removed-Deceased* (SEIRD) para previsão de longo prazo dos casos, recuperados e óbitos somada a resíduos previstos por três modelos ARIMA (ALA'RAJ; MAJDALAWIEH; NIZAMUDDIN, 2021), outro usou o modelo ARIMA para previsão dos casos diários com a previsão dos resíduos pelo modelo MLP (MORAIS; GOMES, 2022). Eles buscam combinar diferentes modelos no intuito de aproveitar as vantagens de cada um e mitigar as respectivas desvantagens

---

para obter um desempenho superior. Nem sempre a combinação de modelos conduz a um resultado superior, pois é preciso que haja uma complementariedade entre suas capacidades e funcionamentos.

Em relação ao objeto da previsão foram identificados trabalhos focados na previsão de casos, enquanto outros fizeram também óbitos e recuperados (COMITO; PIZZUTI, 2022). Quanto ao horizonte de previsão, houve os que fizeram previsão de um passo a frente, e outros fizeram previsões de vários passos.

A Tabela 1 apresenta alguns trabalhos destacados sobre previsão da COVID-19. Observa-se que apesar da diversidade de metodologias e modelos encontrados na literatura, há uma lacuna que não foi preenchida, pois não foram encontrados trabalhos que explorassem padrões locais nas séries temporais.

Todas as abordagens identificadas treinaram os modelos considerando o conjunto de treinamento como um todo. Esse ponto de vista desconsidera que os diferentes ciclos da pandemia apresentam conceitos distintos, que mudam, e se alternam, ao longo do tempo, formando subconjuntos de treinamento. Esta característica possibilita o uso da seleção dinâmica para oferecer adaptabilidade ao sistema preditivo.

O presente trabalho busca, portanto, preencher essa lacuna ao propor um sistema de previsão com seleção dinâmica, que opera com padrões locais para lidar com a mudança de conceito nas séries da COVID-19.

Tabela 1 – Trabalhos sobre Previsão da COVID-19.

Publicação	Grupo	Modelos	Local	Previsão	Início	Métricas	Resultados
(ARDABILI et al., 2020)	ML	LoR, LR, logarithmic, quadratic, cubi, compound, power and exponential regressors, MLP, ANFIS	Alemanha, Itália, Irã, EUA e China	Casos por dia	22/01/20 - 18/03/20	RMSE, correlação	MLP é superior
(ARUNKUMAR et al., 2021)	TS	ARIMA e SARIMA	16 países	Casos, recuperados e óbitos acumulados	22/01/20 - 03/08/20	MSE, MAE, RMSE, MAPE	SARIMA > ARIMA
(CHAURASIA; PAL, 2020)	TS e ML	NB, MA, ES, Holt's linear, Holt-Winters, ARIMA	Índia e global	Mortes	22/01/20 - 28/05/20	RMSE	NB é superior
(DEVARAJ et al., 2021)	TS e DL	ARIMA, LSTM, SLSTM, Prophet	Índia e Chennai	Casos, mortes e recuperação	22/01/20 - 08/05/20	RMSE, MAE, MAPE e R <sup>2</sup>	SLTSM > ARIMA > LSTM
(FANG et al., 2022)	ML	XGB e ARIMA	EUA	Casos 14 dias a frente	13/12/20 - 30/06/21	MAE, RMSE and MAPE	XGBoost > ARIMA
(FIRMINO et al., 2020)	EP,	NCB, SIR, SEIR, SIS, ARIMA e ETS	14 países	Casos por dia	22/01/20 - 26/06/20	MSE, MAPE, Theil, WPOCID	NCB é superior
(HERNANDEZ-MATAMOROS et al., 2020)	TS	ARIMA	145 países	Casos por dia	primeiro caso - 28/05/20	RMSE	RMSE 144,81
(PANDEY et al., 2020)	EP	SEIR, RL	Índia	Casos	30/01/20 - 30/03/20	RMSLE	SEIR > RL
(RIBEIRO et al., 2020)	TS e ML	ARIMA, CUBIST, RF, RIDGE, SVR e EL	Brasil	Casos a um, três e seis dias a frente	24/02/20 - 11/04/20	MAE, sMAPE	SVR e EL são superiores
(TALKHI et al., 2021)	ML	NNETAR, ARIMA, Hybrid, Holt-Winter, BSTS, TBATS, Prophet, MLP, e ELM	Irã	Casos e óbitos 30 dias a frente	20/02/20 - 15/08/20	RMSE, MAE, and MAPE	MLP e Holt-Winter são superiores
(HARROU et al., 2020)	DL	RN, LSTM, BILSTM, GRU, VAE	Itália, Espanha, França, China, Austrália e EUA	Casos	22/01/20 - 17/06/20	RMSE, MAE, MAPE e RMSLE	VAE é superior
(MAALIWI et al., 2021)	TS e DP	ARIMA e S-LSTM	Brasil, EUA, Filipinas e Índia	Casos e óbitos	08/03/20 - 18/06/21	RMSE, MAPE e MBE	Ensemble > ARIMA e S-LSTM
(ALA'RAJ; MAJDA-LAWIEH; NIZAMUDDIN, 2021)	EP e TS	SEIRD e ARIMA	EUA	casos, recuperados e óbitos 2 semanas a frente	22/01/20 - 16/09/20	MAE, MSE e MSLE	Modelo híbrido é superior
(MORAIS; GOMES, 2022)	TS e ML	ARIMA e MLP	Global	Casos diários 7, 14 e 28 dias a frente	22/01/20 - 25/04/21	MASE, SMAPE e R <sup>2</sup>	Modelo híbrido é superior

Fonte: Elaborada pelo autor (2022).

EP: Modelos Epidemiológicos, TS: Modelos Clássicos de Séries Temporais,

ML: Modelos Clássicos de Aprendizado de Máquina e DL: Modelos de Aprendizado Profundo.

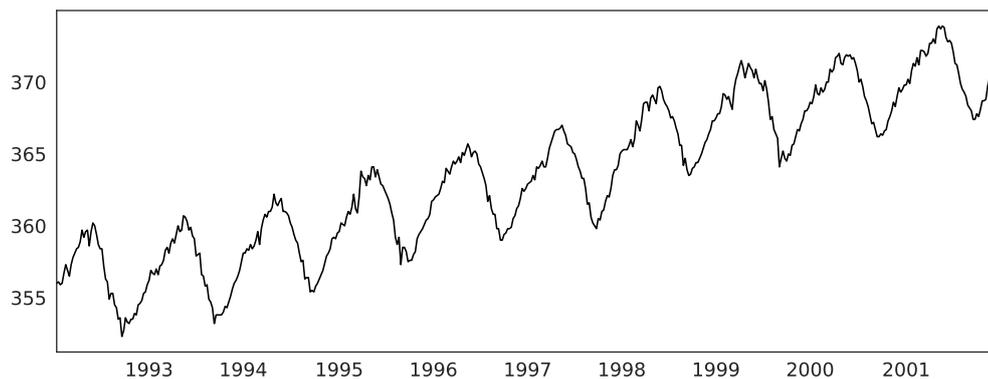
### 3 REFERENCIAL TEÓRICO

Neste Capítulo são apresentados os dois pilares que envolvem a construção do sistema proposto: séries temporais, mudança de conceito e seleção dinâmica.

#### 3.1 SÉRIES TEMPORAIS

Uma série temporal é um conjunto de dados obtidos sequencialmente no tempo. (BOX et al., 2015), e as suas características e modelagem compõem o objeto de estudo da análise e previsão de séries de temporais.

Figura 2 – Série temporal de CO2

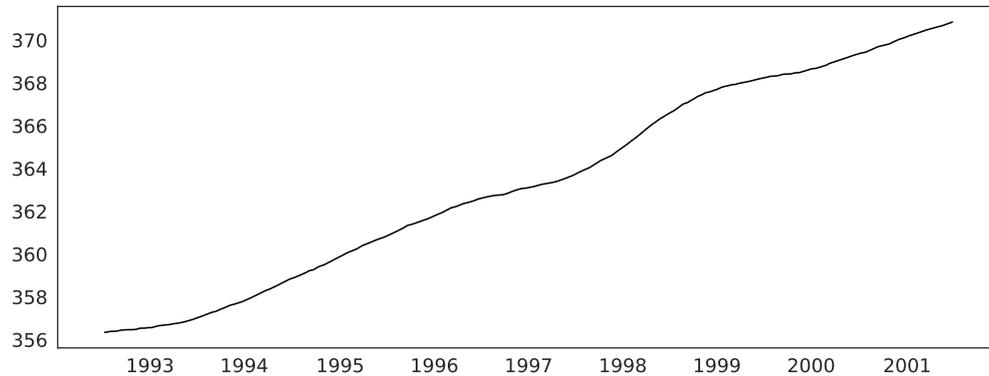


**Fonte:** Elaborada pelo autor (2022). Quantidade de CO2 no ar com dados fornecidos pelo Mauna Loa Observatory, Hawaii, U.S.A..

As séries temporais ( $x_t$ ) podem ser divididas em três componentes: i) a tendência ( $m_t$ ), que representa alterações de valor contínuas e não cíclicas; ii) a sazonalidade ( $s_t$ ), que representa alterações periódicas e cíclicas; iii) resíduo ( $z_t$ ), que representa alterações aleatórias (COWPERTWAIT; METCALFE, 2009).

Em geral, a tendência é mudança sistemática na série, sem periodicidade, que provoca aumentos ou diminuições nos valores da série (COWPERTWAIT; METCALFE, 2009). Na Figura 2, é possível perceber que ao longo do tempo há um aumento no valor médio dos valores, e quando se isola esse componente percebe-se a tendência de crescimento na Figura 3.

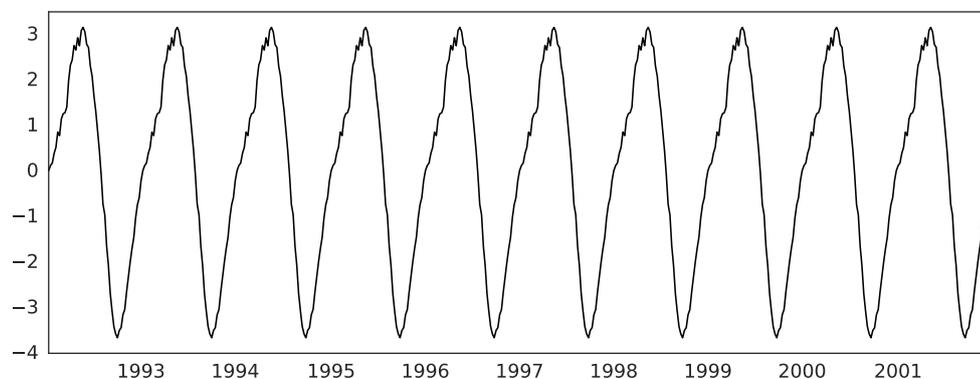
Figura 3 – Tendência da Série temporal de CO2



Fonte: Elaborada pelo autor (2022).

A repetição de um padrão ao longo do ano é a variação sazonal (COWPERTWAIT; METCALFE, 2009). No componente de sazonalidade na Figura 4 nota-se que ao longo do ano há um ciclo de crescimento e declínio.

Figura 4 – Sazonalidade da Série temporal de CO2



Fonte: Elaborada pelo autor (2022).

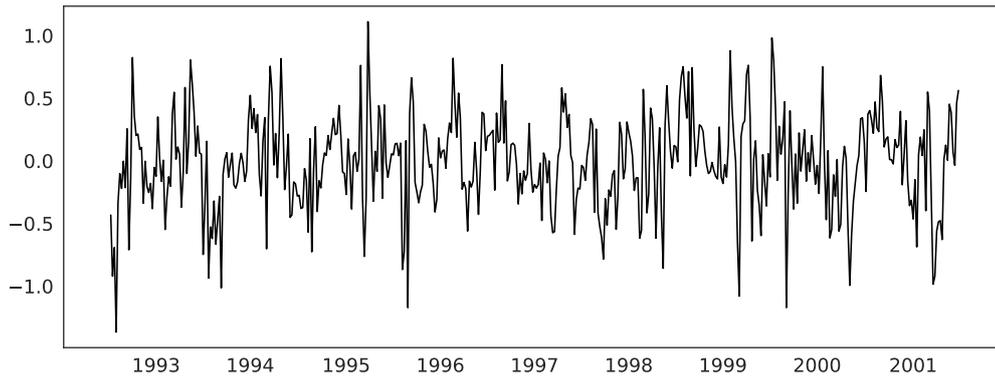
Na decomposição o resíduo corresponde a variação que não é explicada pela tendência nem pela sazonalidade, e possui um comportamento randômico como observado na Figura 5. A decomposição aditiva é definida na seguinte equação:

$$x_t = m_t + s_t + z_t$$

Um diagnóstico relevante nas previsões de séries temporais é quanto aos resíduos da previsão, obtidos pela diferença entre o valor previsto e o real. Quando os erros da previsão são

aleatórios, portanto de forma independente e igual distribuídos com média zero, são designados por ruído branco (COWPERTWAIT; METCALFE, 2009). Este resultado significa que o erro da previsão deve-se ao acaso. Podem ser também que há relações não-lineares, que não foram ajustadas pelo modelo linear.

Figura 5 – Resíduo da Série temporal de CO2



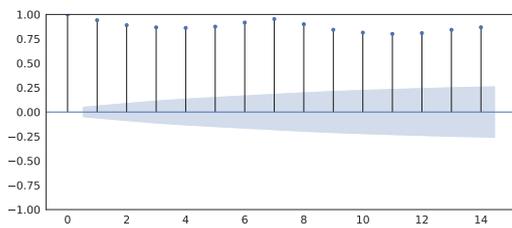
Fonte: Elaborada pelo autor (2022).

As observações adjacentes de uma série temporal podem apresentar relações de dependência significativas entre si configurando uma correlação serial entre os dados, que é aferida pela função de autocorrelação e autocorrelação parcial. A autocorrelação ( $\rho_k$ ) é obtida a partir da função de autocovariância ( $\gamma_k$ ):

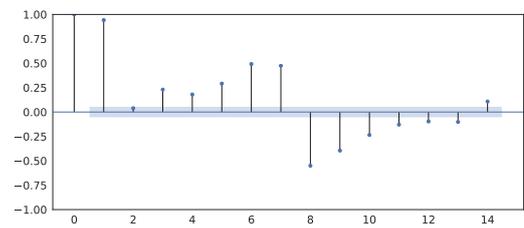
$$\gamma_k = E[(x_t - \mu)(x_{t+k} - \mu)]$$
$$\rho_k = \frac{\gamma_k}{\sigma^2}$$

O estudo da autocorrelação pode ser realizado com o auxílio do correlograma que é uma visualização de dados que permite identificar os *lags* que apresentam correlações significativas. A Figura 6 demonstra um correlograma de autocorrelação e outro de autocorrelação parcial, obtidos a partir dos dados da Figura 2.

Figura 6 – Correlograma



(a) Função de autocorrelação (ACF)



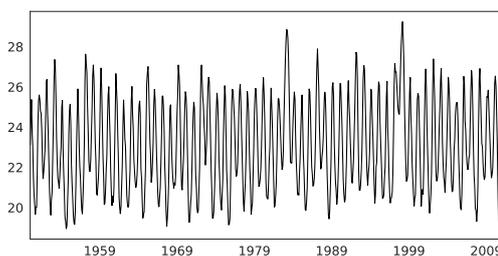
(b) Função de autocorrelação parcial (PACF)

Fonte: Elaborada pelo autor (2022).

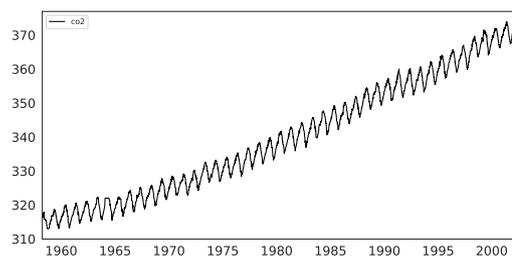
As séries temporais podem ser classificadas em estacionárias e não estacionárias. Quando as propriedades probabilísticas mantêm um equilíbrio que não muda com o tempo, média e variância constante, elas são consideradas estacionárias.

Se a distribuição de dados conjunta de  $x_{t_1}, \dots, x_{t_n}$  é a mesma distribuição conjunta de  $x_{t_1+m}, \dots, x_{t_n+m}$ , e portanto a distribuição não muda após um deslocamento temporal arbitrário, então a série é considerada estritamente estacionária (COWPERTWAIT; METCALFE, 2009). Caso contrário, ela é considerada estacionária de segunda ordem. Caso a média ou variância mudem ao longo do tempo, a série é considerada não estacionária. A Figura 7 apresenta exemplos estacionariedade e não estacionariedade.

Figura 7 – Estacionariedade



(a) Temperatura do Oceano Pacífico - Estacionário



(b) CO2 no ar - Não estacionário

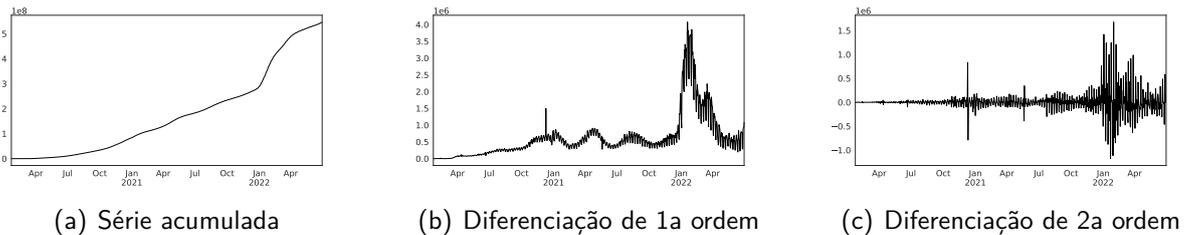
Fonte: Elaborada pelo autor (2022). Esquerda: a temperatura média mensal da superfície do Oceano Pacífico com dados fornecidos pela *National Oceanic and Atmospheric Administration (NOAA)*, direita: quantidade de CO2 no ar com dados fornecidos pelo *Mauna Loa Observatory, Hawaii, U.S.A.*

A diferenciação pode transformar uma série não estacionária em uma série estacionária (COWPERTWAIT; METCALFE, 2009) pela subtração entre os dados da série a ser diferenciada:

$$\nabla x_t = x_t - x_{t-1}$$

Quando a série resultante de uma diferenciação de primeira ordem não é estacionário, é possível realizar sucessivas diferenciações até que se consiga atingir a estacionariedade. A Figura 8 ilustra, a partir da série acumulada de casos de COVID-19, os resultados da diferenciação de primeira e segunda ordem.

Figura 8 – Diferenciações



Fonte: Elaborada pelo autor (2022).

### 3.1.1 Modelos Preditivos

Existem vários modelos preditivos para séries temporais, e dentre eles destaca-se a família de modelos estatísticos ARIMA. Eles são compostos por 2 modelos: i) auto-regressivo (AR) e ii) média móvel (MA).

O modelo auto-regressivo é definido pela seguinte equação:

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + w_t$$

onde  $w_t$  é o ruído branco e  $\alpha_i$  são os parâmetros do modelo para  $\alpha_p \neq 0$  para um modelo de ordem ( $p$ ) (COWPERTWAIT; METCALFE, 2009).

O modelo média móvel é definido pela seguinte equação:

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_q x_{t-q} + w_t$$

onde  $w_t$  é o ruído branco com média zero e variância  $\sigma^2$  e  $\beta_i$  são os parâmetros do modelo (COWPERTWAIT; METCALFE, 2009).

A combinação desses dois modelos resulta no modelo ARMA (auto-regressivo média móvel) de ordem ( $p, q$ ). Pode ser acrescida uma operação de diferenciação de ordem  $d$  para transformar a série em estacionária a ser ajustada pelo modelo ARMA. A inclusão desse operador (I) gera o modelo ARIMA de ordem ( $p, d, q$ ).

O modelo ARIMA pode incorporar parâmetros para lidar com a sazonalidade gerando o *seasonal* ARIMA (SARIMA) de ordem  $(p,d,q)(P,D,Q)$  (COWPERTWAIT; METCALFE, 2009), definido pela seguinte equação:

$$\Theta_{\rho}(B^S)\theta_p(B)(1 - B^S)^D(1 - B)^d x_t = \phi_Q(B^S)\phi_q(B)w_t$$

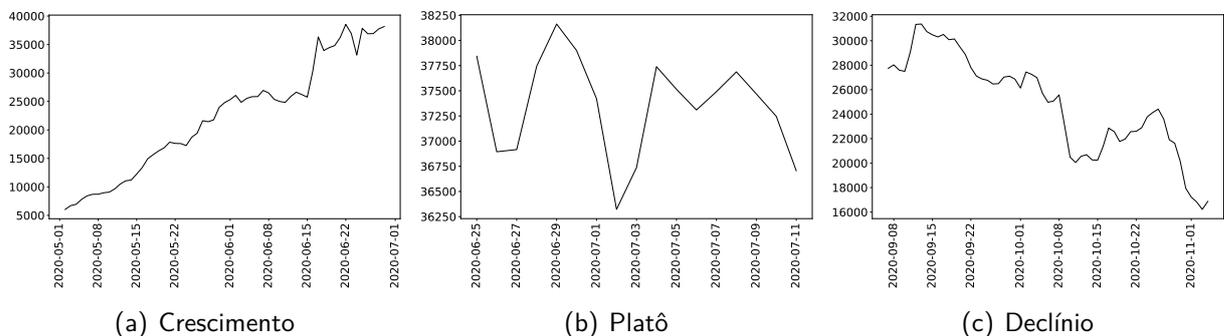
onde  $\Theta_{\rho}$ ,  $\theta_p$ ,  $\phi_Q$  e  $\phi_q$  são polinômios de ordens P, p, Q, and q, respectivamente, e  $Bx_t = x_{t-1}$ .

A definição do modelo pode ser realizada usando a metodologia de Box e Jenkins (BOX et al., 2015) que envolve as seguintes atividades: i) escolha do modelo candidato, ii) estimação dos parâmetros e iii) avaliação da adequação do modelo. Após a realização destas atividades, caso os resultados obtidos não sejam aceitáveis reinicia-se o ciclo de atividades, caso sejam aceitáveis a definição do modelo é concluída.

### 3.2 MUDANÇA DE CONCEITO

No contexto da pandemia, as diferentes fases nos ciclos da COVID-19 podem configurar conceitos distintos, e apresentarem comportamentos díspares, com três padrões principais: crescimento, platô e declínio (FIRMINO et al., 2020). O crescimento tem tendência positiva, platô tem tendência neutra e o declínio uma tendência negativa.

Figura 9 – Padrões da tendência da COVID-19 no Brasil



Fonte: Elaborada pelo autor (2022)

Essas mudanças de conceito dificultam a tarefa de previsão por um único modelo (MINKU; YAO, 2012). Por exemplo, se o modelo está ajustado com um viés para uma tendência de declínio e há uma mudança para uma tendência de crescimento declínio, aquele sistema de

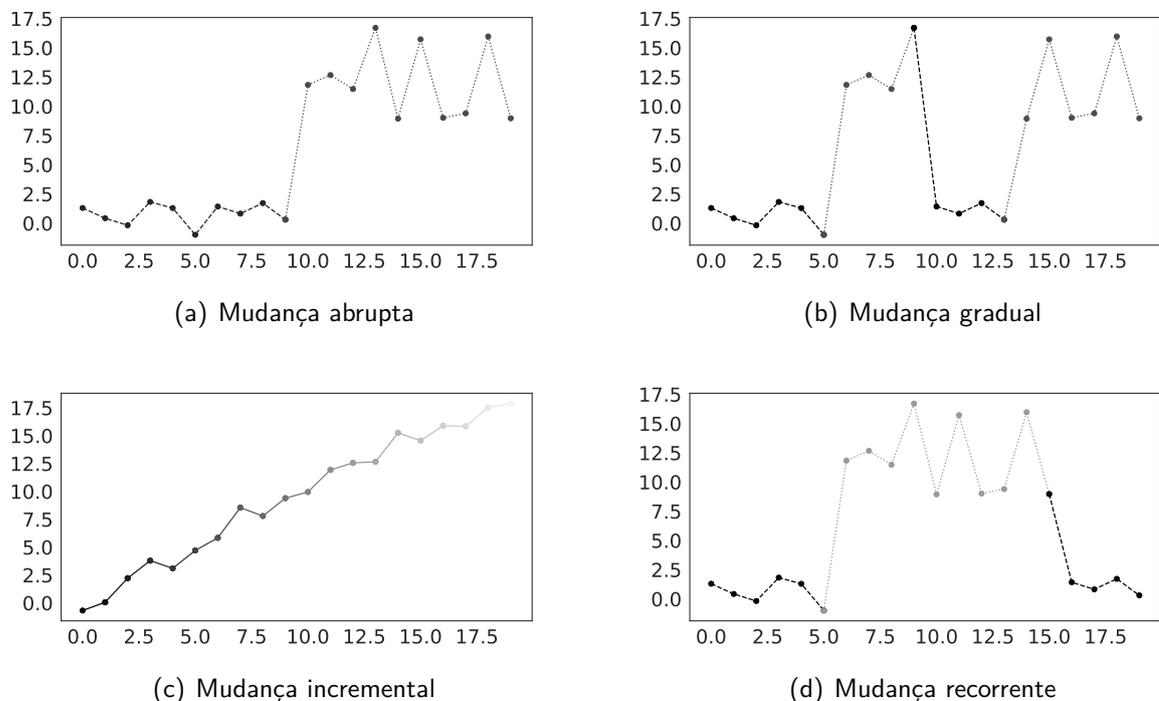
previsão terá uma degradação no seu desempenho, por não ter sido capaz de reconhecer e se adaptar a outra tendência (RASHED; HIRATA, 2021). Assim, é interessante decompor o problema da previsão dos casos da COVID-19 em problemas da previsão nos diferentes conceitos.

Um conceito pode ser definido formalmente como a distribuição de probabilidade conjunta do problema  $p_t(X, y) = p_t(X|y)$  (WEBB et al., 2016).

Em séries temporais não estacionárias, cujas média e variância se alteram ao longo do tempo, é possível identificar a ocorrência de mudanças de conceito ao longo da série. Isso implica uma alteração na distribuição dos dados, o que pode ser prejudicial para modelos preditivos, que não aprenderam ou não reconhecem essa modificação.

A literatura apresenta diversas classificações para essas mudanças de conceito, de acordo com diferentes critérios de análise (WEBB et al., 2016). Baseando-se no modo como elas ocorrem é possível classificá-las como mudança abrupta, gradual, incremental e recorrente (ŽLIOBAITĖ, 2010).

Figura 10 – Tipos de Mudança de Conceito.



**Fonte:** Elaborada pelo autor (2022). Exemplos com dados sintéticos.

A mudança abrupta, Figura 11(a), caracteriza-se por um tempo curto de transição para o novo conceito. Um exemplo seria a transição repentina do padrão de declínio para o surto

de crescimento de casos da variante ômicron (REN et al., 2022), pelo aumento a taxa de transmissão da doença.

A mudança gradual, Figura 11(b), caracteriza-se por um tempo extenso de transição para o novo conceito, em que há uma coexistência entre o padrão vigente e o sucessor. Um exemplo seria a transição de padrão de declínio para o crescimento pela alteração na predominância da variante alfa para a delta (FLORENSA et al., 2022), que possuíam taxas de transmissão distintas.

A mudança incremental, Figura 11(c), caracteriza-se por um tempo extenso de transição para o novo conceito, em que não se identifica o conceito vigente nem o novo conceito até que se complete a mudança. Um exemplo seria a transição do número de óbitos de um padrão de crescimento para o declínio na medida em que a vacinação aumentava (CHRISTIE et al., 2021), pela redução na taxa de mortalidade.

A mudança recorrente, Figura 11(d), caracteriza-se por alternância em padrões. Um exemplo seria as transições da tendência de crescimento para declínio motivadas pela vigência, ou não, de *lockdowns* (HSIANG et al., 2020), pelas alterações na taxa de transmissão.

A identificação da mudança de conceito é muito importante para tomada de decisão e adaptação ao novo padrão da série temporal. Essa tarefa é realizada por meio de algoritmos de detecção de mudança que são capazes de avaliar se a distribuição de dados foi alterada, e portanto se está diante de um novo padrão (ŽLIOBAITÉ, 2010). Isso não se confunde com a detecção de *outliers* que não implica mudança no conceito, mas a ocorrência de anomalias em pontos de dados na série descorrelacionadas com o restante da série e sem correspondência com a distribuição dos dados.

As soluções para lidar com a mudança de conceito podem ser classificadas em implícitas e explícitas (CAVALCANTE; MINKU; OLIVEIRA, 2016). As soluções implícitas propõem um atualização periódica dos modelos preditivos regularmente, independente da ocorrência de mudança de conceito, mas caso ocorra alguma mudança, os modelos já estarão atualizados e adaptados. Uma desvantagem dessa abordagem é constante atualização e consumo de recursos, mesmo que não haja mudança, e portanto um novo treinamento do modelo não seja necessário. Além do que nessa abordagem não se mede nem analisa a mudança de conceito, o que no caso da pandemia é uma informação tão ou mais importante que a própria previsão.

As soluções explícitas propõem o monitoramento de estatísticas das séries para identificar a ocorrência da mudança. Uma abordagem comum é monitorar a distribuição dos resíduos. Se houver uma alteração, isso significa que o modelo não está mais ajustado aos dados, e portanto é preciso uma adequação ao novo conceito. Outra abordagem é monitorar a distribuição das

variáveis preditoras. Se é detectada uma alteração na distribuição das variáveis, isso pode implicar uma mudança dos dados /emph(data drift), quando não há alteração na distribuição de probabilidades conjunta, mas pode ser mudança de conceito /emph(concept drift), quando há mudanças na distribuição.

A utilização de múltiplos preditores tem sido mostrado uma solução para o problema da mudança de conceito, especialmente sob o enfoque da diversidade dos modelos como proposto por (MINKU; YAO, 2012).

### 3.3 SELEÇÃO DINÂMICA

Diante de problemas como a previsão dos casos da COVID-19, muitas vezes é preciso utilizar mais de um modelo para resolver o problema. De acordo com o teorema do *no free lunch* (WOLPERT, 1996), não há modelo que seja capaz de resolver todos os problemas, por isso é preciso identificar a solução adequada para cada caso.

A utilização de vários modelos pode ser feita por meio de sistemas de múltiplos preditores (SMP) (SILVA; NETO; CAVALCANTI, 2021). Esta solução envolve três etapas principais: geração, seleção e combinação (SERGIO; LIMA; LUDERMIR, 2016).

Inicialmente, é realizada a geração dos preditores por meio do treinamento dos modelos buscando otimizar a diversidade (MINKU; WHITE; YAO, 2010) entre eles para que sejam complementares, e consigam gerar coletivamente um conjunto de previsões com menor erro do que os modelos únicos.

Após a geração, é preciso selecionar o melhor ou os melhores candidatos para resolver o problema, o que pode ser realizado estática ou dinamicamente. A seleção estática ocorre quando o método de seleção do preditor, ou conjunto de preditores, é definido *a priori* com base no melhor desempenho no conjunto de validação, por exemplo atribuindo pesos diferentes para cada preditor. Depois dessa definição, a seleção não sofre influência do padrão de teste.

A seleção dinâmica é executada em função de cada padrão de teste que é submetido ao sistema de previsão, seguindo a regra estabelecida para selecionar os modelos, portanto a seleção é influenciada dinamicamente pelos padrões de teste.

Um parâmetro muito utilizado pela literatura para avaliação da qualidade da seleção dinâmica é a comparação com o Oracle, que consiste na melhor opção de seleção para determinado padrão de teste, representando o desempenho que poderia ser atingida, caso houvesse uma seleção perfeita (CRUZ; SABOURIN; CAVALCANTI, 2017).

Por fim, caso sejam selecionados vários preditores deve ser feita a combinação das suas previsões para a obtenção da previsão final. Isso pode ser feito por meio de estatísticas, como a média, mediana ou moda (KOURENTZES; BARROW; CRONE, 2014) . É possível otimizar e atribuir pesos distintos para as previsões com o uso algoritmos de inteligência de enxames (LIMA et al., 2021). Há outras estratégias como o uso de um modelo secundário, por exemplo uma rede neural, para efetuar a combinação das previsões (RIBEIRO et al., 2020).

A literatura apresenta diversos modelos construídos com a aplicação da seleção dinâmica propondo inovações nas diferentes etapas que envolvem os SMP. Dentre os quais destacam-se modelos híbridos que operam com previsão de resíduos (OLIVEIRA; SILVA; NETO, 2021), definição de região de competência (SILVA; NETO; CAVALCANTI, 2021; DAI; SONG; YAO, 2019) e combinação de preditores (SERGIO; LIMA; LUDERMIR, 2016). Ainda que haja uma grande variedade aplicações nestes trabalhos, não foram identificadas publicações que utilizassem a seleção dinâmica para lidar com o problema da previsão de casos e óbitos do novo coronavírus.

A seleção dinâmica afigura-se como um método interessante para fornecer adaptabilidade ao sistema preditivo, o que é essencial para resolver o problema da mudança de conceito nas séries pandêmicas.

Após decompor a série em várias regiões de competência, ou conceitos como se tem utilizado, a seleção dinâmica responde a pergunta de qual modelo utilizar, e integrar a solução, completando a estratégia dividir para conquistar.

O método de seleção dinâmica aplicado às series da COVID-19 deve ser invariante à diferentes escalas, por isso o uso da distância entre o padrão treinamento e padrão de teste não foi utilizado (SILVA et al., 2020).

## 4 SISTEMA PROPOSTO

O sistema CODYS pode ser dividido em duas etapas principais: treinamento e teste. Na etapa de treinamento, o conjunto de treinamento é dividido em séries menores representando padrões distintos por meio da detecção de conceito. Depois disso, os modelos de previsão são ajustados a essas subséries. Na etapa de teste para cada padrão de teste é apresentada anteriormente uma região de competência, o sistema prevê os valores para essa região e, com base no menor erro, seleciona os melhores K modelos. Por fim, realiza a previsão do padrão de teste para cada um dos K modelos, e depois é feita uma combinação dessas K previsões.

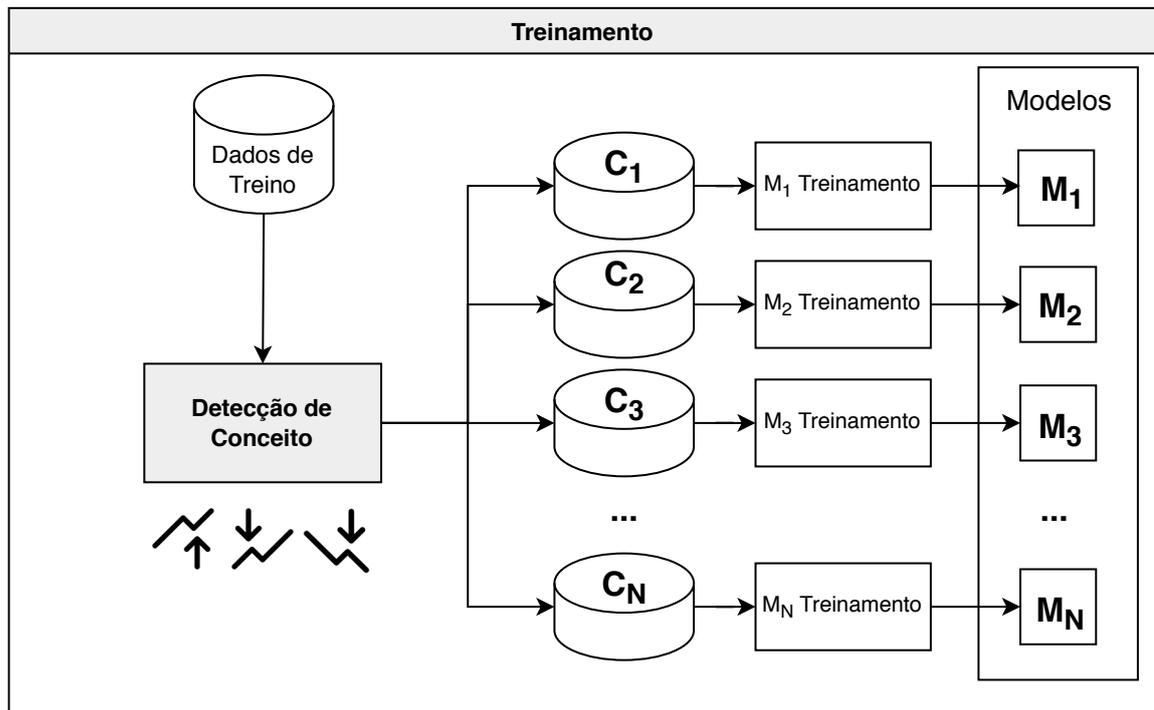
### 4.1 ETAPA DE TREINAMENTO

Durante a etapa de treinamento (veja a Fig. 11), o sistema proposto recebe o conjunto de treinamento de uma série temporal  $Z_t$ , como dados de entrada. Em seguida, a detecção da mudança de conceito é aplicada, a fim de dividir o conjunto de dados em conceitos distintos.

Existem vários tipos de detectores de mudança. A diversidade de tipos de mudança motivou a criação desta variedade de detectores, pois o desempenho de cada um deles é afetado pela forma como mudança de conceito ocorre. O estudo comparativo de (GONÇALVES et al., 2014) revela as particularidades e performances de vários detectores em diferentes conjuntos de dados. Dentre eles, encontra-se o *Adaptive Windowing (ADWIN)* que se destacou nos testes realizados pela pequena quantidade de falsos positivos e acurácia na segregação entre os diferentes conceitos, e por isso foi a opção adotada neste trabalho.

*ADWIN* utiliza janelas deslizantes de tamanho variável que são redefinidas dinamicamente, de acordo com a distribuição dos dados destas janelas. Um subconjunto de dados da série temporal é selecionado e sub dividido em duas janelas ( $W_0, W_1$ ) que são usadas para identificar se houve alguma mudança (BIFET; GAVALDA, 2007). É realizado um teste de hipótese comparando a média de  $W_0$  e  $W_1$  para determinar se os dados correspondem a mesma distribuição. A mudança de conceito é detectada, quando a hipótese nula de igualdade das distribuições é rejeitada. Quando isso acontece,  $W_0$  é substituída por  $W_1$  e uma nova  $W_1$  é inicializada. Os intervalos de confiança são construídos, de acordo com o hiperparâmetro  $\alpha \in (0, 1)$ , os quais são usados nos testes.

Figura 11 – Etapa de treinamento.



**Fonte:** Elaborada pelo autor (2022). Primeiro, os dados são submetidos à análise de detecção de desvio para dividir as séries em N conceitos diferentes. Depois, as divisões são usadas para treinar um especialista em previsão em cada conceito para compor um pool de modelos.

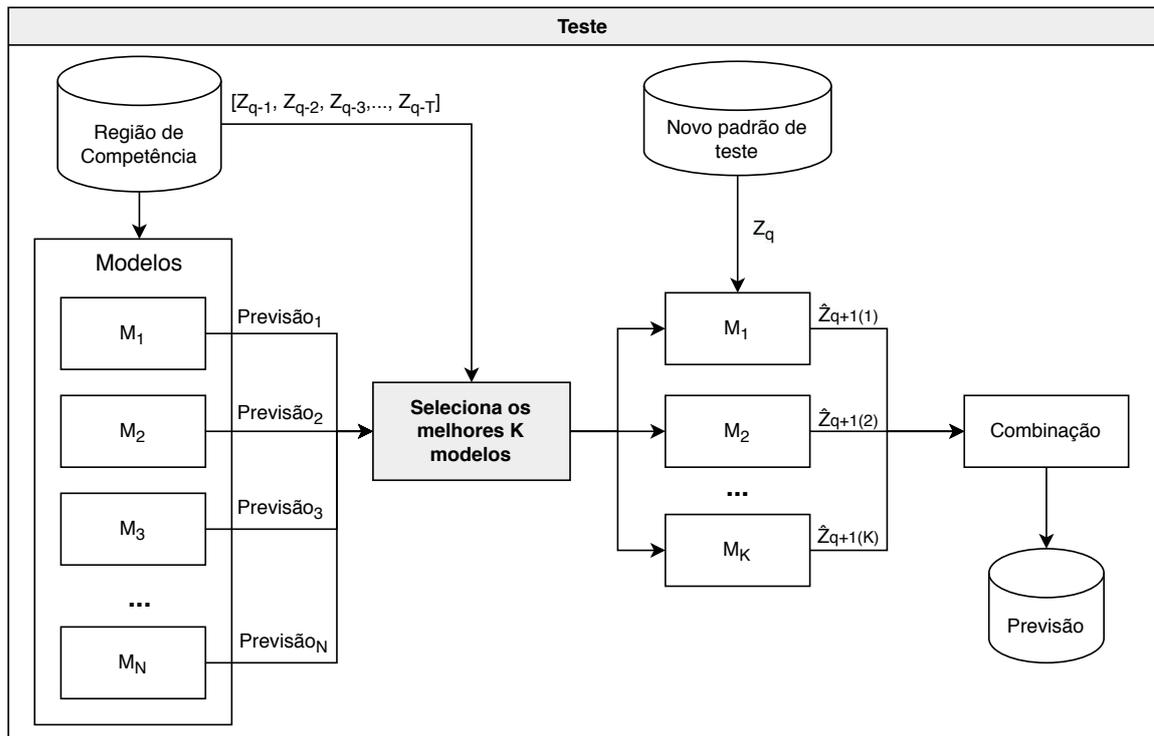
Cada subsérie será utilizada para treinar um, ou mais modelos, naquele conceito específico. A escolha de qual tipo de modelo treinar é muito importante, pois a série de treino pode ser muito curta, e assim o modelo escolhido deve conseguir ser treinado com poucos pontos de dados.

Ao final do treinamento, é obtido um conjunto de modelos especialistas ajustados aos diferentes padrões da séries.

## 4.2 ETAPA DE TESTE

Após a etapa de treinamento, há um conjunto de modelos especialistas nas fases de crescimento, platô e declínio. O próximo desafio é selecionar quais e quantos modelos serão usados para prever cada instância, o que é realizado dinamicamente (veja a Fig. 12).

Figura 12 – Etapa de teste.



**Fonte:** Elaborada pelo autor (2022). Os modelos recebem dados de entrada para prever as próximas  $T$  etapas de tempo. As previsões resultantes são pontuadas em relação ao padrão de teste. Os modelos são classificados de acordo com sua competência para prever o padrão de teste. A seleção leva  $K$  modelos para prever o(s) próximo(s) passo(s) à frente. As previsões dos modelos são combinadas com a previsão do sistema de saída.

Inicialmente é apresentada uma região de competência  $(Z_q - 1, Z_q - 2, Z_q - 3, \dots, Z_{q-T})$  construída a partir de  $T$  lags em relação ao padrão de teste  $Z_q$  (SILVA; NETO; CAVALCANTI, 2021). A utilização destes lags mais próximos tem a finalidade de criar uma região de competência com o mesmo conceito do padrão de teste a ser previsto, assumindo que os dados vizinhos possuem conceitos iguais.

Após obter as previsões, a métrica de erro é calculada e os modelos são ordenados do menor ao maior erro. A partir do ranking, são selecionados os  $K$  melhores modelos.

A seleção dos  $K$  melhores preditores recebe os dados de entrada do padrão de teste  $(Z_q)$  e calcula as  $K$  previsões, as quais, ao final, são combinadas pela média ou mediana.

## 5 METODOLOGIA DOS EXPERIMENTOS

Neste Capítulo é apresentado o protocolo experimental adotado na pesquisa.

### 5.1 DESCRIÇÃO DOS DADOS

O conjunto de dados usados nos experimentos é composto por séries temporais dos casos diários da COVID-19 em um grupo de países selecionados considerando o critério geográfico. América: Brasil e Estados Unidos; Europa: Inglaterra, Alemanha, França e Itália; África: África do Sul, e Ásia: Rússia, Índia e Japão. O período estudado foi de 22/01/2020 a 31/12/2021. Estas séries apresentam diferentes padrões de tendência e sazonalidade. Os dados foram fornecidos pela John Hopkins University (DONG et al., 2022).

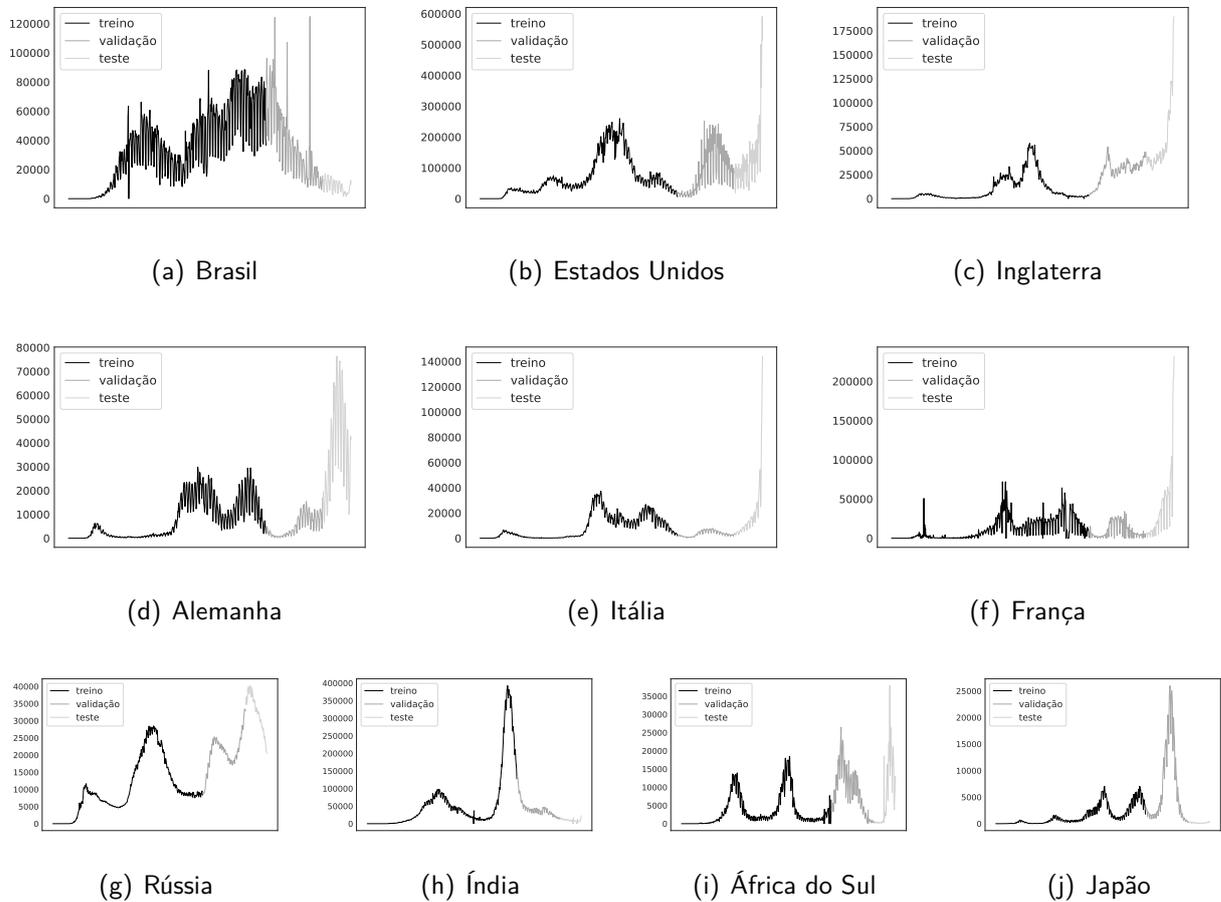
As séries de cada país foram divididas em três sub series: treinamento de 22/01/2020 até 31/05/2021 (496 pontos, 70%), validação de 01/06/2021 até 20/10/2021 (142 pontos, 20%) e teste de 22/10/2021 até 31/12/2021 (71 pontos, 10%).

Tabela 2 – Estatísticas dos conjuntos de dados

País	Média	Desvio	25%	50%	75%	Máximo	Soma
Brasil	31.402	25.183	10.285	27.380	47.671	124.878	22.264.355
Estados Unidos	76.629	73.823	26.747	53.445	109.693	590.666	54.329.920
Inglaterra	18.126	23.184	1.940	7.108	31.095	189.836	12.851.348
Alemanha	10.002	13.296	947	4.764	13.908	76.414	7.091.365
Itália	8.597	12.067	1.147	4.458	13.382	144.255	6.095.460
França	13.859	21.270	1.094	6.469	20.314	232.200	9.826.325
Rússia	14.499	10.271	6.569	10.955	22.562	40.210	10.280.129
Índia	48.901	74.632	9.765	25.467	50.848	392.488	34.671.162
África do Sul	4.853	5.765	772	2.112	7.983	37.875	3.440.981
Japão	2.434	4.261	199	877	2.666	25.992	1.725.937

**Fonte:** Elaborada pelo autor (2022).

Figura 13 – Separação dos Dados.



**Fonte:** Elaborada pelo autor (2022). Treinamento em preto, validação em cinza escuro e teste em cinza claro. Ordem de apresentação por critério geográfico.

Os conjuntos de dados foram usados da seguinte forma:

Tabela 3 – Relação entre tarefa da previsão por tipo de modelo e conjunto de dados usados.

Tarefa/Modelo	Estatísticos	Aprendizado de Máquina	CODYS
Detecção de Conceito	-	-	Treinamento e Validação
Busca de hiperparâmetros	Treinamento e Validação	Validação	Treinamento e Validação
Ajuste aos dados	Treinamento e Validação	Treinamento e Validação	Treinamento e Validação
Avaliação	Teste	Teste	Teste

**Fonte:** Elaborada pelo autor (2022).

Para os modelos de *machine learning* (SVR, ELM, MLP e LSTM), a busca de hiperparâmetros realizou-se com base no conjunto de validação, a partir dos modelos inicialmente treinados com os dados de treinamento. Os melhores hiperparâmetros foram usados para especificar os modelos que foram ajustados aos dados usando, nesse segundo momento, os dados de trei-

namento e validação, conjuntamente. Após essa tarefa, a avaliação dos modelos foi realizada com os dados de teste. Um particularidade para os modelos de ML foi que o ajuste dos dados e avaliação repetiu-se dez vezes para cada modelo, de forma a evitar o viés nos resultados, por conta da inicialização aleatória dos parâmetros deste modelos.

Para os modelos estatísticos (ARIMA, SARIMA), o treinamento inicial, a busca de hiperparâmetros e o ajuste aos dados usaram conjuntamente os dados de treinamento e validação. Após essa tarefa, a avaliação dos modelos foi realizada com os dados de teste. Não houve repetição para estes modelos.

Para o CODYS, houve uma tarefa inicial e diferente dos demais de detecção dos modelos usando o conjunto de treinamento e validação. A partir das subdivisões resultantes dessa tarefa, foi realizada a busca por hiperparâmetros e ajuste aos dados para cada subdivisão. Ao final, a avaliação dos modelos foi realizada com os dados de teste.

## 5.2 MODELOS DE PREVISÃO

Os modelos subjacentes aos sistema proposta foram variações do Seasonal Autoregressive Integrated Moving Average (SARIMA). A pequena quantidade de dados é um desafio para a previsão dos casos da COVID-19, pois os modelos não conseguem aprender os padrões com poucas informações, e assim ficam mais sujeitos ao subajuste ou sobreajuste.

O *baseline* foi determinado pelo método *randow walk*, que consiste em usar o último valor observado como previsão para o próximo passo.

Os modelos de comparação (OLIVEIRA; SILVA; NETO, 2021) foram selecionados entre modelos estatísticos: ARIMA, SARIMA e modelos de Machine Learning: Support Vector Machines Regressor (SVR), Multi Layer Perceptron (MLP), Extreme Learning Machines (ELM), Long Short Term Memory (LSTM) e Extreme Gradient Boosting (XGB).

Além disso, os modelos citados anteriormente foram utilizados para compor *ensembles* com combinações das previsões pela média e mediana.

A otimização de hiperparâmetros dos modelos foi executada antes do treinamento dos modelos para atingir desempenhos confiáveis. O ajuste de modelos pode ser uma tarefa muito desafiadora devido a grandes números de opções e a sensibilidade deles.

Utilizou-se um a seguinte grade de busca (OLIVEIRA; SILVA; NETO, 2021).

Tabela 4 – Hiperparâmetros dos Modelos.

Modelo	Parâmetros	Valores
ARIMA	p, q	[0,7]
	d	1
	critério	aic
SARIMA	p, q	[0,7]
	d	1
	$p_s, q_s$	[0,7]
	critério	aic
SVR	<i>kernel</i>	linear, rbf, sigmoid
	<i>gamma</i>	$10^{-3}, 10^{-2}, 10^{-1}, 10^0$
	C	$10^{-1}, 10^0, 10^1, 10^2, 10^3, 10^4$
	<i>epsilon</i>	$10^{-3}, 10^{-2}, 10^{-1}$
	tolerância	$10^{-4}, 10^{-3}, 10^{-2}$
	<i>lags</i>	14
	neurônios	2, 5, 10, 15, 20, 50, 100
MLP	ativação	tanh
	otimizador	Adam
	taxa de aprendizado	$10^{-4}, 10^{-3}, 10^{-2}$
	função de perda	Huber
	<i>lags</i>	14
	tamanho do lote	32
	épocas	100
ELM	neurônios	5, 10, 100, 500
	ativação	tanh
	<i>lags</i>	14
LSTM	neurônios	5, 10, 100, 500
	ativação	tanh
	otimizador	Adam
	taxa de aprendizado	$10^{-4}, 10^{-3}, 10^{-2}$
	função de perda	Huber
	<i>lags</i>	14
	tamanho do lote	32
XGB	épocas	100
	<i>booster</i>	gbtree, gblinear, dart
	<i>lambda</i>	$[10^{-8}, 10^{-0}]$
	<i>alpha</i>	$[10^{-8}, 10^{-0}]$
	profundidade máxima	[1,10]

Fonte: Elaborada pelo autor (2022).

No caso do ARIMA e SARIMA, a busca foi realizada usando o auto-arima (SMITH et al., 2017–) para escolher a melhor ordem pelo teste dos termos autoregressivo (p), média móvel (q) e diferenciação (d) para encontrar o Critério de Informação de Akaike (AIC) no modo *stepwise*. O número máximo de lags foi definido em sete, baseado no estudo das funções de autocorrelação (ACF) e autocorrelação parcial (PACF). Esta otimização foi realizada para os modelos únicos e o sistema proposto.

No caso dos modelos de aprendizado de máquina, a busca é feita pelo método Tree-structured Parzen Estimator (TPE)(AKIBA et al., 2019) que recebe as opções do espaço de busca e seguindo os caminhos mais próximos para atingir o ótimo. O número de *lags* foi fixado em 14 baseando-se também no estudo do ACF e PACF, entretanto no caso dos modelos de aprendizado de máquina a quantidade de 7 *lags* foi insuficiente e como havia correlações relevantes até o *lag* 14, esse valor foi usado. Isso levou a um desempenho médio superior se comparado a quantidade menores ou maiores de *lags*.

### 5.3 PRÉ-PROCESSAMENTO, DETECÇÃO DE CONCEITO E SELEÇÃO DINÂMICA

Antes de disponibilizar os dados para os modelos, alguns procedimentos de pré-processamento foram aplicados: remoção de *outliers* e normalização.

O conjunto de dados original é uma série diária de casos acumulados que são fornecidos pelos órgãos de saúde pública dos países (DONG et al., 2022), mas este tipo de agregação não possibilitava diagnosticar as mudanças de conceitos, por isso as séries foram diferenciadas para refletindo a variação diária dos casos.

Depois da diferenciação, notou-se que as séries acumuladas estavam escondendo alguns pontos de dados errôneos. Alguns deles resultavam em números negativos, e outros implicavam em enormes aumentos que eram incompatíveis com as séries. Para evitar tais erros, os números negativos foram substituídos por zeros, e as séries de treinamento foram limitadas ao nonagésimo nono percentil.

Normalização utilizada transformou os valores para o intervalo entre -1 e 1. A série normalizada foi usada pelos modelos de aprendizado de máquina para treinamento e previsão, exceto no caso da SVR que usou dados de entrada normalizada e a saída os dados na escala original, devido a problemas no treinamento usando a saída normalizada. Os modelos estatísticos utilizaram os dados na escala original, e os cálculos das métricas também.

A detecção de conceito foi realizada por meio do algoritmo ADWIN, com uma janela mínima

de 30 dias e alfa de 0,005.

A seleção dinâmica foi realizada usando janelas temporais de sete passos prévios ao alvo, que representa a semana prévia. Para avaliar a competência dos modelos na região foi usada o Mean Squared Error (MSE), que se provou ser uma boa opção dentre outras métricas.

MSE é uma métrica para medir o desempenho da previsão pelo erro médio quadrático. O valor é calculado pela diferença média quadrada entre as previsões e os respectivos valores reais (ARUNKUMAR et al., 2021), quanto menor o valor, melhor é o modelo. A métrica MSE é definida por meio da seguinte equação, onde  $y$  é valor real, e  $\hat{y}$  é o valor previsto,  $n$  é quantidade de valores:

$$\text{MSE}(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2.$$

A quantidade  $K$  de modelos selecionados foi igual a um, de forma que somente o modelo com menor erro foi escolhido. Em testes realizados com  $K > 1$  não foi possível identificar um valor de  $K$  que resultasse desempenho médio superior ao obtido com  $K = 1$ , portanto os resultados apresentados só consideram a seleção de um modelo.

Em termos de pós-processamento, as previsões com valores negativos foram substituídas por zero, pois a série temporal é composta por números naturais.

#### 5.4 MÉTRICAS DE DESEMPENHO

Um conjunto de métricas foi usado para oferecer uma avaliação mais ampla da performance dos modelos considerando métricas de previsão comuns como Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) and Coefficient of Determination ( $R^2$ ).

MAE é uma métrica utilizada para medir o desempenho da previsão pelo erro médio absoluto. O valor é calculado pela diferença absoluta média entre as previsões e os respectivos valores reais (ARUNKUMAR et al., 2021), quanto menor o valor, melhor é o modelo. O MAE é definido pela seguinte equação, onde  $y$  é o valor real, e  $\hat{y}$  é o valor previsto,  $n$  é a quantidade de valores:

$$\text{MAE}(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i|.$$

RMSE é uma métrica para medir o desempenho da previsão pela raiz quadrada do erro médio quadrático. O valor é calculado pela raiz quadrada da diferença média quadrada entre

as previsões e os respectivos valores reais (CHAURASIA; PAL, 2020), quanto menor o valor, melhor é o modelo. Diferencia-se do MAE por penalizar previsões que são mais distantes dos valores reais. A métrica RMSE é definida por meio da seguinte equação, onde  $y$  é o valor real, e  $\hat{y}$  é o valor previsto,  $n$  é a quantidade de valores:

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2}.$$

MAPE é uma métrica utilizada para medir o desempenho da previsão pelo erro médio absoluto percentual. O valor é calculado pela diferença absoluta média entre as previsões e os respectivos valores reais divididas pelos valores reais (ARUNKUMAR et al., 2021), quanto menor o valor, melhor é o modelo. Uma vantagem do seu uso é padronizar uma escala comum [0,100] para dados com escalas diferentes, por isso foi a métrica utilizada nos resultados. O MAPE é definido pela seguinte equação, onde  $y$  é o valor real, e  $\hat{y}$  é o valor previsto,  $n$  é a quantidade de valores:

$$\text{MAPE}(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} \frac{|y_i - \hat{y}_i|}{\max(\epsilon, |y_i|)}.$$

$R^2$  é uma métrica utilizada para medir o desempenho da previsão pela variância do erro (DEVARAJ et al., 2021). O valor é calculado pela razão entre soma dos quadrados dos resíduos e Soma Total dos Quadrados. O resultado determina o quanto da variância é explicada pelo modelo [0,1], quanto maior o valor, melhor é o modelo. O  $R^2$  é definido pela seguinte equação, onde  $y$  é o valor real, e  $\hat{y}$  é o valor previsto,  $n$  é a quantidade de valores:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}.$$

## 5.5 ESPECIFICAÇÃO DE SOFTWARE E HARDWARE

Todos experimentos foram realizados em um máquina usando Ubuntu 20.04 LTS 64-bit, INTEL i5-10600KF 4.10GHz x 12 CPU, NVIDIA GeForce GTX1650 4GB GPU e 16GB RAM.

Os modelos LSTM e MLP foram construídos usando Tensorflow 2.9, ELM com o Pytorch 1.11, SVM com o scikit-learn 1.0.2, e ARIMA e SARIMA foi com statsmodels 0.13.2.

Os resultados foram registrados e gerenciados usando a plataforma MIFlow.

## 6 RESULTADOS DOS EXPERIMENTOS

Neste Capítulo os resultados dos experimentos são apresentados. Inicialmente é abordada a detecção de conceito nas diferentes séries, buscando avaliar se os conceitos identificados são estatisticamente diferentes. Em seguida, o desempenho do sistema proposto é comparado com outras soluções. Na sequência há uma comparação entre o modelo estático, SARIMA, e o dinâmico, CODYS. Posteriormente, é exibido um comparativo entre a seleção dinâmica e o oracle. Por fim, os tempos de treinamento e teste dos diferentes modelos são comparados.

### 6.1 ANÁLISE DOS CONCEITOS DETECTADOS

O detector de mudança de conceito foi capaz de segmentar a série em regiões de competência para treinar os modelos especialistas. Ainda que o *ADWIN* se restrinja a avaliar a diferença entre as médias nas diferentes janelas temporais, isso foi suficiente para determinar regiões com diferentes tendências e sazonalidades. A Figura 14 mostra as séries e os pontos de mudança detectados.

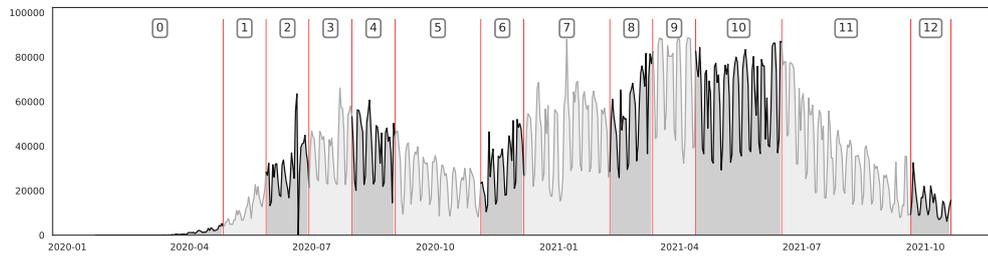
Cada série apresenta um comportamento único, o que resultou em quantidade diferente de conceitos. Alguns países vivenciaram mais picos e variância que outros. Também é possível perceber que o número de conceitos é aproximadamente três vezes o número de ciclos, o que é esperado se consideradas as três principais fases do ciclo: crescimento, estabilidade e declínio.

Considerando o ponto de mudança, essa é a uma tarefa mais difícil, porque até mesmo para um especialista humano é desafiador definir exatamente quando um conceito termina e o próximo começa, o que não é claro em mudanças graduais ou incrementais. Entretanto, a detecção de mudança foi capaz de encontrar pontos razoáveis para separar as séries e delimitar os conceitos. Para avaliar isto, testes estatísticos de Mann–Whitney U foram realizados para analisar se as janelas adjacentes pertencem a mesma distribuição.

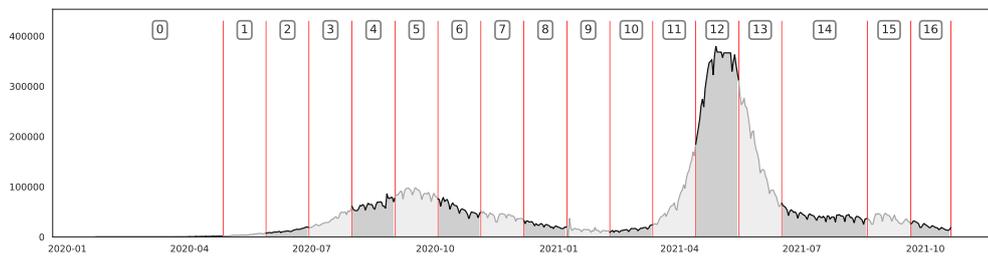
Os resultados dos testes indicam que na maioria dos casos as janelas adjacentes pertencem as distribuições distintas, o que é destacado na última linha da Tabela 5. Isto prova que a técnica aplicada apresenta resultados válidos.

A detecção de mudança pode, portanto, ser um método útil para identificar conceitos nas séries temporais, quando é inviável, ou difícil estabelecer manualmente, o ponto de mudança, e pode ser aplicada em sistemas adaptativos e não supervisionados.

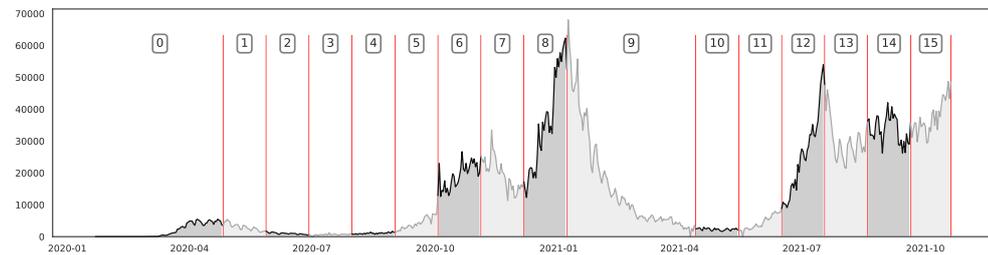
Figura 14 – Detecção dos Conceitos dos Países.



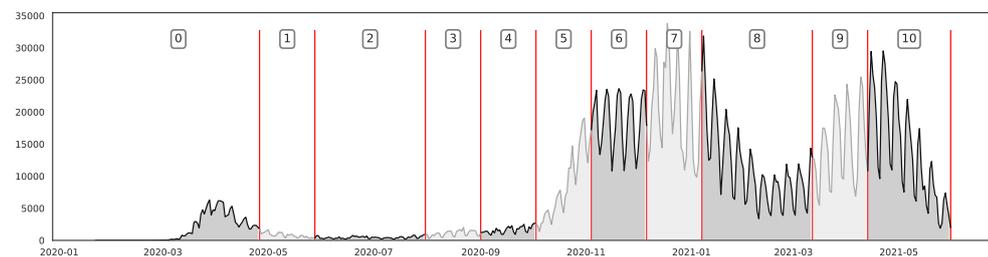
(a) Brasil



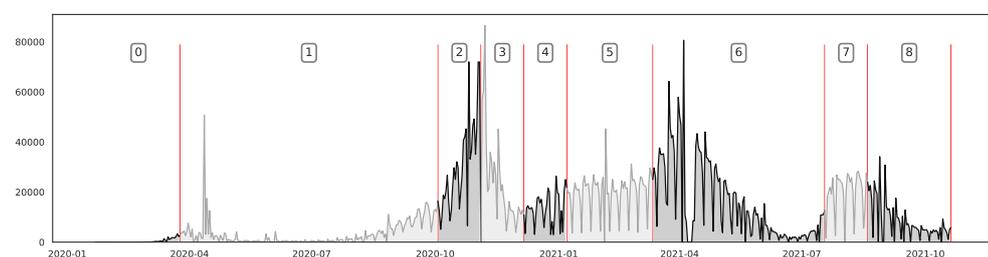
(b) Estados Unidos



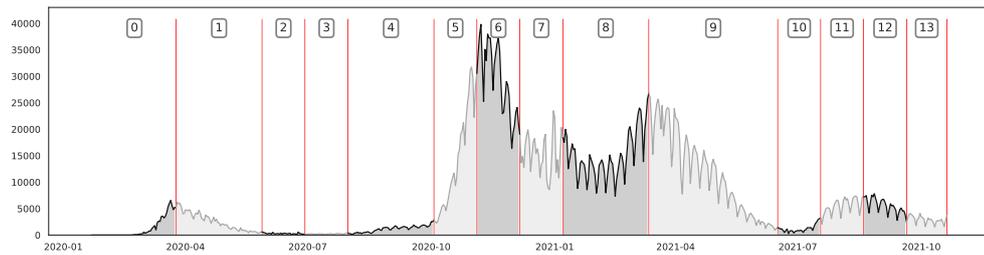
(c) Inglaterra



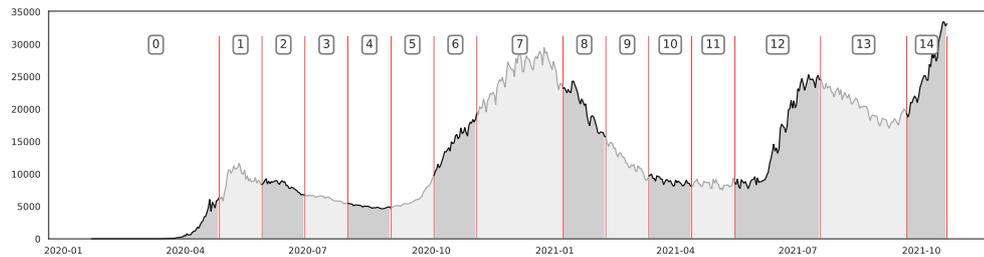
(d) Alemanha



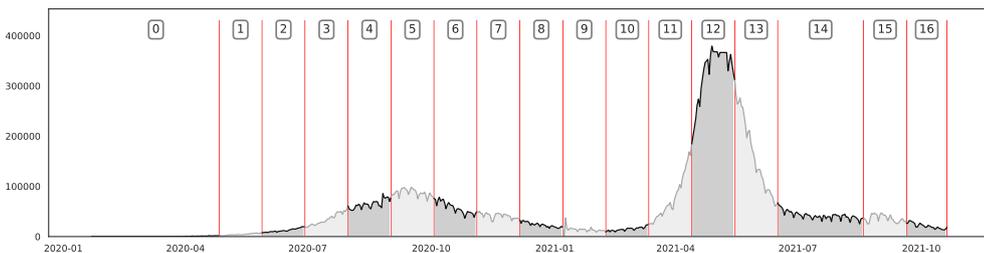
(e) França



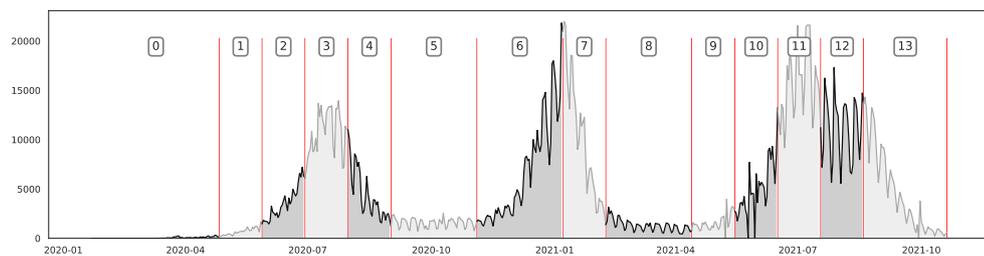
(a) Itália



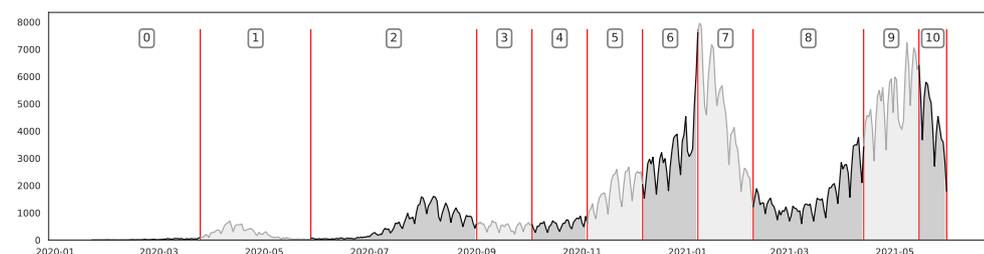
(b) Rússia



(c) Índia



(d) África do Sul



(e) Japão

**Fonte:** Elaborada pelo autor (2022). Cada cor representa um conceito e as linhas vermelhas marcam a separação entre eles.

Tabela 5 – Resultados do teste de Mann–Whitney U comparando as distribuições de duas janelas adjacentes.

Janelas Adjacentes	Brasil	EUA	Inglaterra	Alemanha	França	Itália	Rússia	Índia	África do Sul	Japão
0 - 1	≠	≠	≠	=	≠	≠	≠	≠	≠	≠
1 - 2	≠	≠	≠	≠	≠	≠	≠	≠	≠	≠
2 - 3	≠	=	≠	≠	=	≠	≠	≠	≠	=
3 - 4	=	≠	≠	≠	=	≠	≠	≠	≠	=
4 - 5	≠	≠	≠	≠	≠	≠	≠	≠	≠	≠
5 - 6	≠	≠	≠	≠	≠	≠	≠	≠	≠	≠
6 - 7	≠	≠	=	=	≠	≠	≠	≠	≠	≠
7 - 8	≠	=	≠	≠	≠	=	≠	≠	≠	≠
8 - 9	≠	≠	≠	≠		≠	≠	≠	=	≠
9 - 10	≠	≠	≠	=		≠	≠	=	≠	≠
10 - 11	≠	≠	≠			≠	≠	≠	≠	
11 - 12	≠	≠	≠			=	≠	≠	≠	
12 - 13		≠	=			≠	≠	≠	≠	
13 - 14			≠				≠	≠		
14 - 15			≠					≠		
15 - 16								≠		
Razão	1/12	2/13	2/15	3/10	2/8	2/13	0/13	1/16	1/13	2/10

**Fonte:** Elaborada pelo autor (2022). As janelas tem distribuições diferentes (denotado por  $\neq$ ) se o  $p$ -valor é menor que 0,05. Por outro lado, os conceitos iguais tem a mesma distribuição (denotado por  $=$ ). A razão mostra o número de janelas adjacentes com a mesma distribuição sobre o total de janelas criadas pelo método ADWIN.

## 6.2 COMPARATIVO DE DESEMPENHO ENTRE MODELOS

A análise comparativa demonstra que o desempenho do sistema proposto, mensurado pelo MAPE, é superior na previsão de casos diários da COVID-19, conforme Tabela 6.

Tabela 6 – Comparação em termos de MAPE do CODYS com modelos concorrentes.

Modelo	Brasil	EUA	Inglaterra	Alemanha	França	Italia	Rússia	India	África do Sul	Japão
RW	35.797	55.484	9.32	31.373	64.911	22.404	2.123	19.977	60.702	32.229
ARIMA	30.195	28.887	9.303	<u>15.512</u>	64.838	16.736	2.1	20.169	50.775	46.075
SARIMA	35.447	28.556	8.461	<b>15.377</b>	55.865	16.969	2.205	23.185	47.515	<u>25.934</u>
SVR	<u>29.919</u>	<b>25.120</b>	8.089	55.84	55.291	<u>15.495</u>	<u>1.853</u>	<u>14.203</u>	<u>43.299</u>	29.005
ELM	60.744	32.523	15.459	86.893	53.552	19.966	7.59	31.697	93.831	99.62
MLP	87.708	<u>26.640</u>	18.946	30.552	<b>51.963</b>	23.786	16.742	208.911	228.848	355.61
LSTM	95.321	37.805	23.059	44.092	66.049	28.689	22.419	249.275	62.533	282.183
XGB	67.569	49.835	23.767	49.356	54.891	21.911	16.442	22.061	51.459	180.525
ENS MÉDIA	50.309	30.31	<u>14.862</u>	31.07	55.576	19.44	7.95	71.292	75.307	32.032
ENS MEDIANA	57.015	28.266	17.057	36.171	55.863	17.898	5.566	21.273	53.233	33.501
CODYS	<b>23.637</b>	36.636	<b>6.974</b>	15.527	<u>52.701</u>	<b>15.481</b>	<b>1.852</b>	<b>12.226</b>	<b>36.447</b>	<b>19.689</b>
ORACLE	8.735	13.633	2.836	3.86	18.983	9.456	0.889	6.794	20.272	6.873

**Fonte:** Elaborada pelo autor (2022). A última linha apresenta o Oracle do CODYS, o menor erro possível a partir das suas previsões. Negrito é o melhor modelo, e sublinhado é o segundo melhor.

Tabela 7 – Diferença percentual entre os modelos.

Modelo	Brasil	EUA	Inglaterra	Alemanha	França	Italia	Rússia	India	Africa do Sul	Japão
RW	51.44	120.88	33.64	104.03	24.92	44.72	14.63	63.4	66.55	63.69
ARIMA	27.74	15	33.4	<u>0.88</u>	24.78	8.11	13.39	64.97	39.31	134.01
SARIMA	49.96	13.68	21.32	<b>0.00</b>	7.51	9.61	19.06	89.64	30.37	<u>31.72</u>
SVR	<u>26.58</u>	<b>0.00</b>	15.99	263.14	6.4	<u>0.09</u>	<u>0.05</u>	<u>16.17</u>	<u>18.8</u>	47.32
ELM	156.99	29.47	121.67	465.08	3.06	28.97	309.83	159.26	157.45	405.97
MLP	271.06	<u>6.05</u>	171.67	98.69	<b>0.00</b>	53.65	804	1608.74	527.89	1706.14
LSTM	303.27	50.5	230.64	186.74	27.11	85.32	1110.53	1938.89	71.57	1333.20
XGB	185.86	98.39	240.79	220.97	5.63	41.53	787.8	80.44	41.19	816.88
ENS MÉDIA	112.84	20.66	<u>113.11</u>	102.06	6.95	25.57	329.27	483.12	106.62	62.69
ENS MEDIANA	141.21	12.52	144.58	135.23	7.51	15.61	200.54	74	46.06	70.15
CODYS	<b>0.00</b>	45.84	<b>0.00</b>	0.98	<u>1.42</u>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>

**Fonte:** Elaborada pelo autor (2022). O melhor modelo, em negrito, e os demais. O segundo melhor está sublinhado.

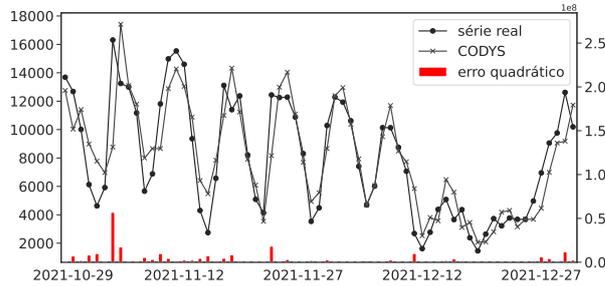
Além de superar os outros modelos na maioria dos casos, foi capaz de obter uma significativa melhoria considerando o segundo melhor modelo da comparação. A Tabela 7 demonstra que o sistema proposto foi 31% e 26% superior do que o melhor concorrente, nas séries do Japão e Brasil, respectivamente.

A Figura 15 demonstra que os melhores modelos foram capazes de prever a tendência de crescimento ou declínio, porém o CODYS foi mais competente em prever após mudanças de conceito. Isto pode ser visto no fim da série do Brasil, França, Índia, Itália e EUA quando o modelo de comparação falhou em seguir a explosão ou desabamento dos casos

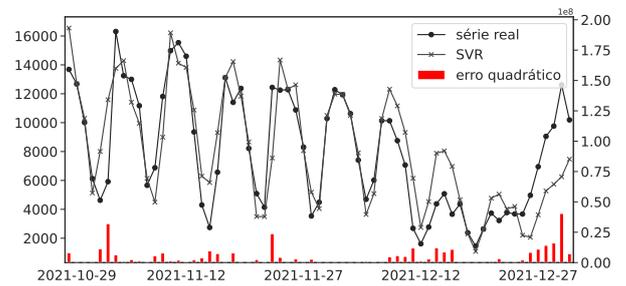
A análise estatística (DEMŠAR, 2006) pelo teste pareado de *Wilcoxon* entre o CODYS e demais modelos, o único modelo que não rejeita a hipótese nula para um  $\alpha$  de 0.05 é a SVR, conforme Figura 16, provando a superioridade do desempenho do CODYS em relação aos concorrentes.

A comparação das distribuições dos resultados apresentada na Figura 17 demonstra que o ranking, a partir do MAPE, do CODYS é superior aos demais, pois o modelo foi o melhor ou segundo melhor, exceto em dois países, portanto foi mais preciso e consistente que os demais.

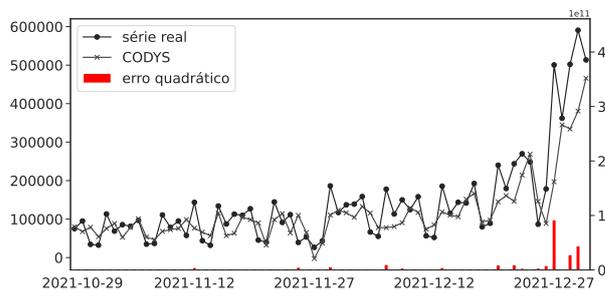
Figura 15 – Previsão de CODYS e melhor modelo concorrente.



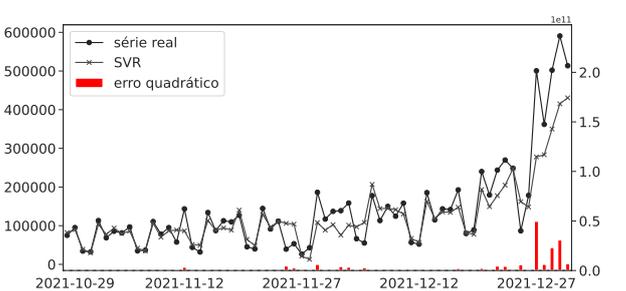
(a) Previsão do CODYS para Brasil



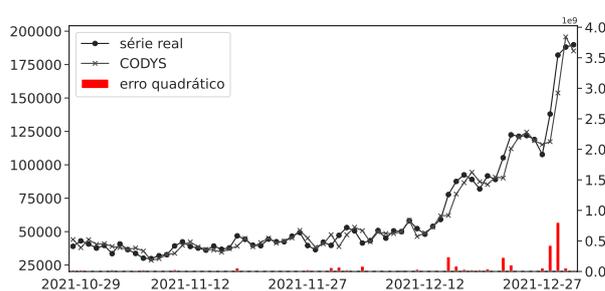
(b) Previsão da SVR para Brasil



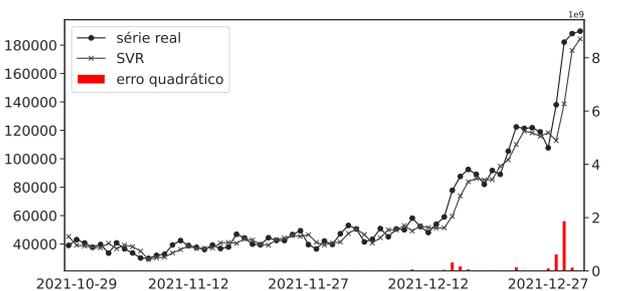
(c) Previsão do CODYS para EUA



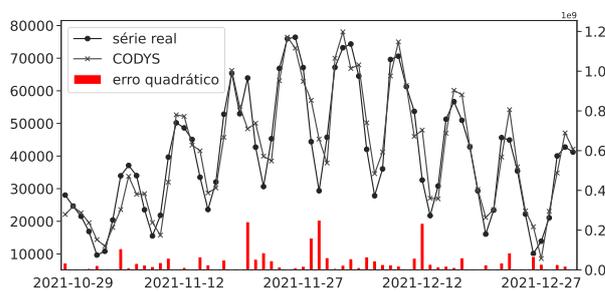
(d) Previsão da SVR para EUA



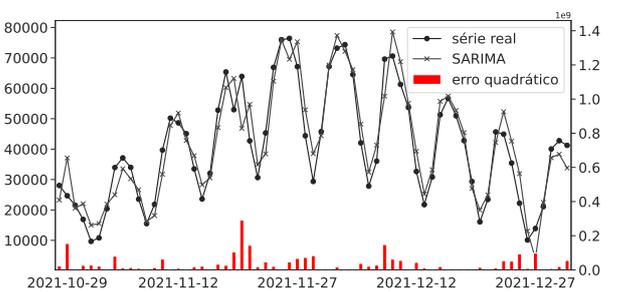
(e) Previsão do CODYS para Inglaterra



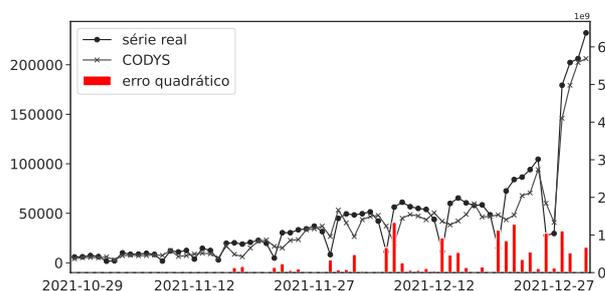
(f) Previsão da SVR para Inglaterra



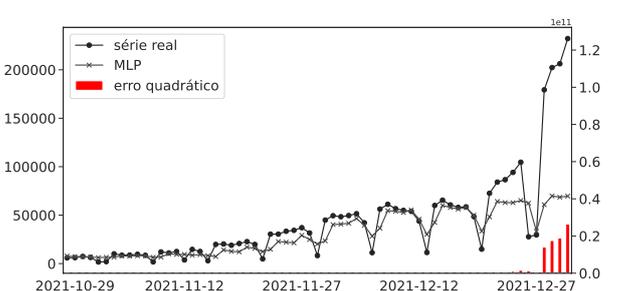
(g) Previsão do CODYS para Alemanha



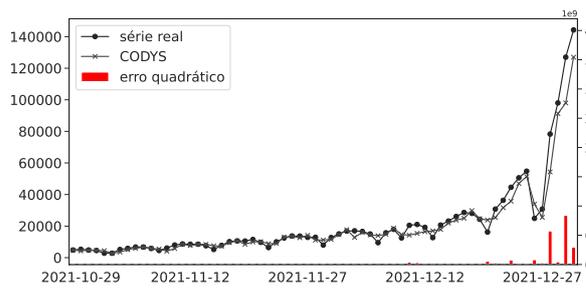
(h) Previsão do SARIMA para Alemanha



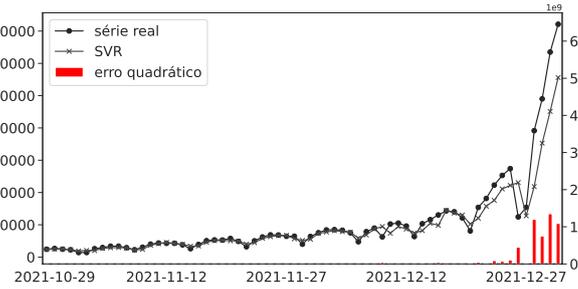
(i) Previsão do CODYS para França



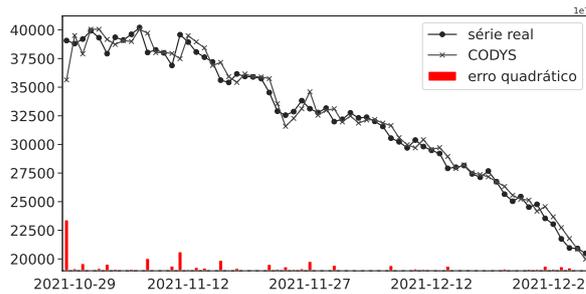
(j) Previsão da MLP para França



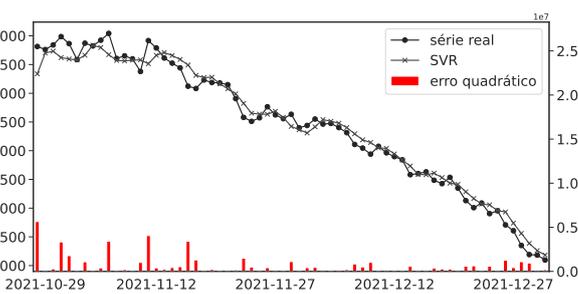
(a) Previsão do CODYS para Itália



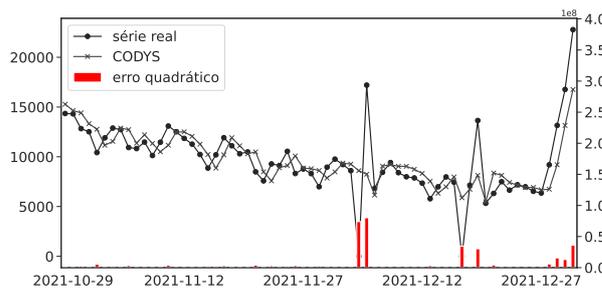
(b) Previsão da SVR para Itália



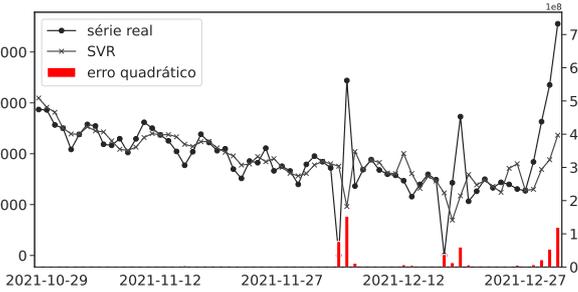
(c) Previsão do CODYS para Rússia



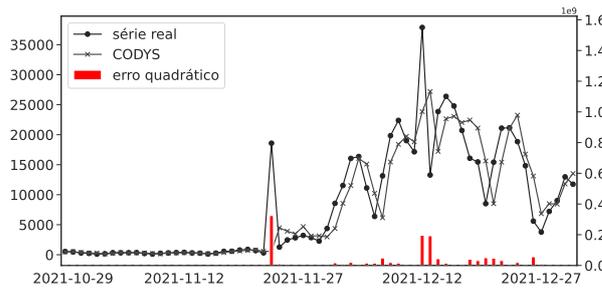
(d) Previsão da SVR para Rússia



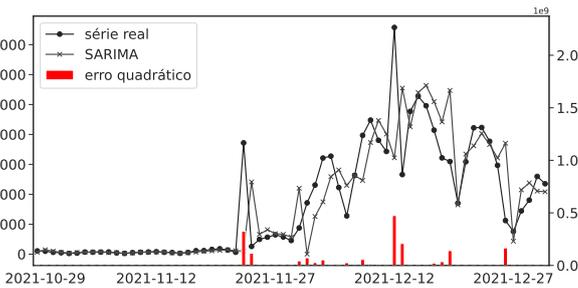
(e) Previsão do CODYS para Índia



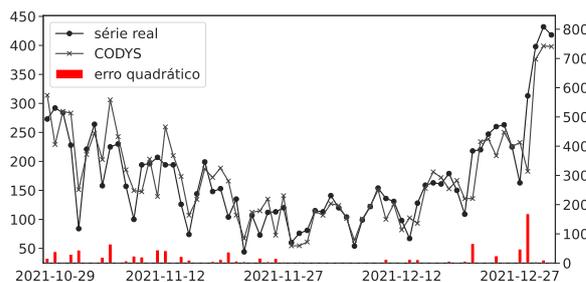
(f) Previsão da SVR para Índia



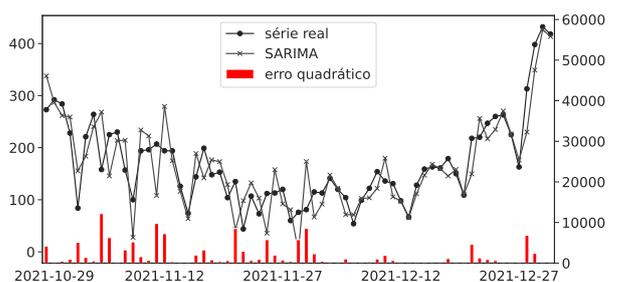
(g) Previsão do CODYS para África do Sul



(h) Previsão do SARIMA para África do Sul



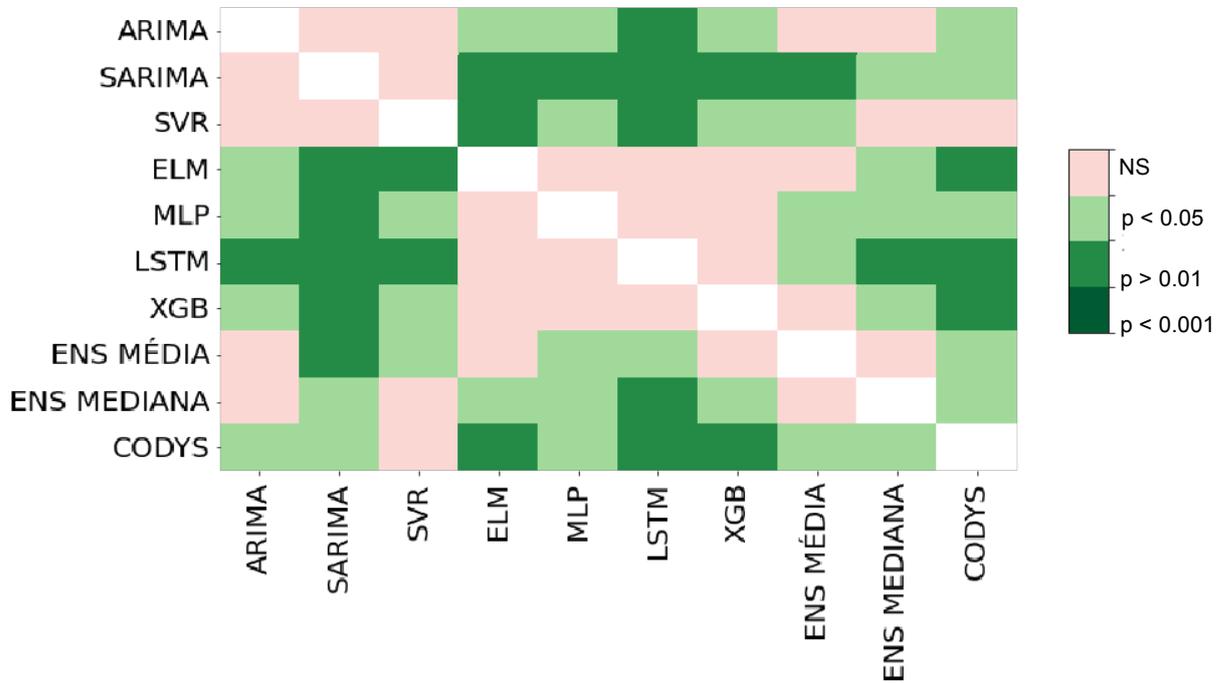
(i) Previsão do CODYS para Japão



(j) Previsão do SARIMA para Japão

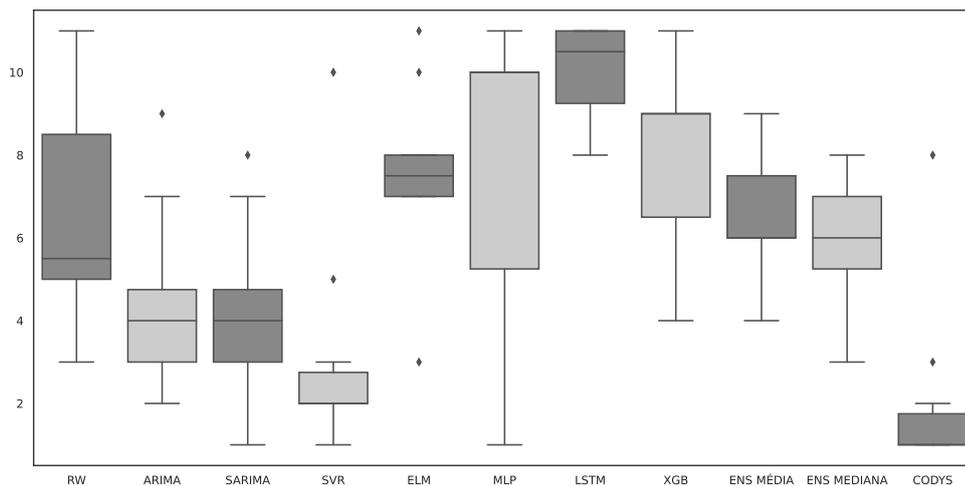
Fonte: Elaborada pelo autor (2022). As barras vermelhas representam o erro quadrático para cada ponto.

Figura 16 – Teste de Wilcoxon.



Fonte: Elaborada pelo autor (2022). NS: não há significância estatística da diferença.

Figura 17 – Boxplot do ranking dos modelos.



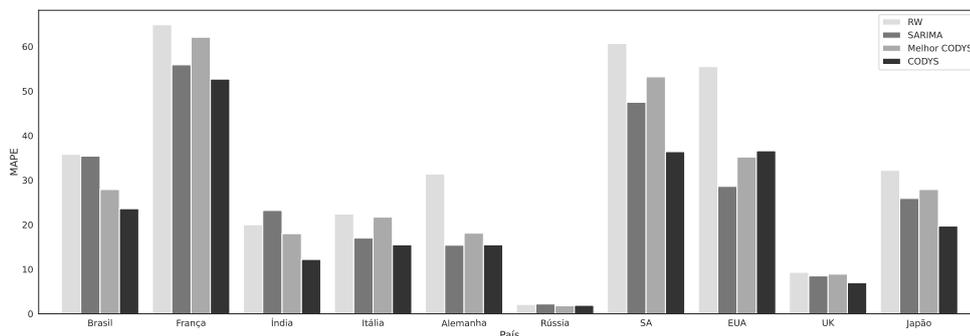
Fonte: Elaborada pelo autor (2022).

### 6.3 COMPARATIVO ENTRE O MODELO ESTÁTICO E DINÂMICO

A comparação entre o modelo estático e o dinâmico revela os ganhos desta abordagem, conforme se observa na Figura 18:

1. As previsões do modelo SARIMA obtém um MAPE inferior ao obtido pelo *baseline* RW, exceto na Índia e Rússia;
2. O melhor modelo do *pool*<sup>1</sup> do CODYS é capaz de superar o *baseline* RW na Índia e Rússia, e supera o SARIMA no Brasil, entretanto perde nos demais;
3. O CODYS com o auxílio da seleção dinâmica consegue selecionar modelos mais adequados a partir do *pool* para previsão, e assim supera o *baseline*, SARIMA e o melhor modelo do *pool* em todos os países, exceto nos EUA e na Alemanha.

Figura 18 – Comparativo entre modelo estático e dinâmico



**Fonte:** Elaborada pelo autor (2022). Na métrica MAPE quanto menor o valor, melhor.

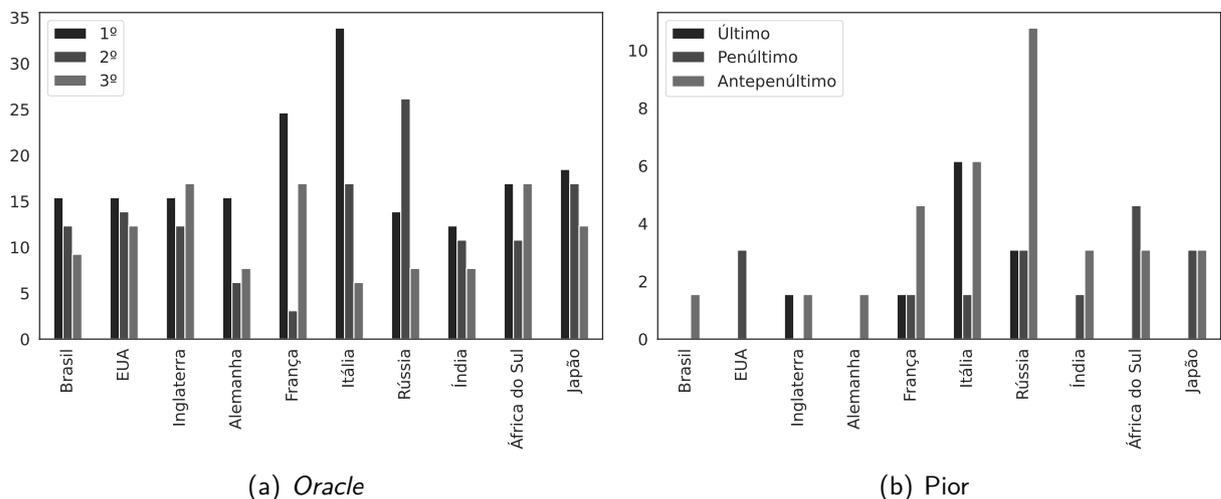
Infere-se, portanto, que a partir de um modelo superior ao *baseline*, é possível gerar um conjunto de modelos que supera o modelo único, entretanto é necessária uma estratégia de seleção dinâmica, pois o melhor modelo individualmente do *pool* não é suficiente para obter um desempenho superior.

<sup>1</sup> Conjunto de modelos

## 6.4 DESEMPENHO DA SELEÇÃO DINÂMICA EM RELAÇÃO AO ORACLE

O sucesso do sistema depende da seleção dinâmica e quão bem ela é capaz selecionar o modelo mais adequado para a previsão do passo temporal alvo. Isto implica selecionar boas opções e evitar as ruins. A seleção conseguiu uma alta porcentagem entre os três melhores modelos, e baixa porcentagem entre os três piores. Nos casos da Alemanha, África do Sul, Brasil, Estados Unidos, Índia e Japão o pior modelo nunca foi selecionado, isto explica a ausência de algumas barras na Figura 19.

Figura 19 – Percentual de seleção de cada modelo do pool.

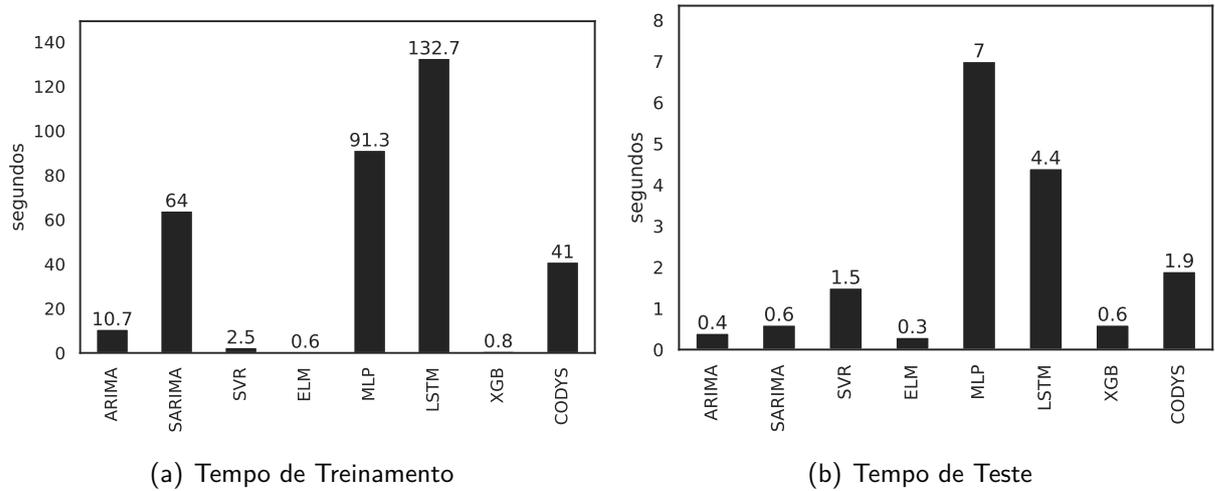


**Fonte:** Elaborada pelo autor (2022). Considerou-se a classificação de desempenho de cada modelo.

## 6.5 ANÁLISE DO TEMPO COMPUTACIONAL

Considerando o tempo de treinamento do CODYS, a sua performance foi melhor do que as redes neurais (MLP e LSTM) e o SARIMA. Os múltiplos modelos do CODYS foram otimizados e ajustados mais rápido do o modelo singular, o que pode ser devido à necessidade de processar os dados de todo o conjunto de treinamento pelo auto-arma pelo SARIMA, enquanto nos modelos especialistas do CODYS as séries são menores. Outros modelos de machine learning tiveram (SVR, ELM, XGB) tiveram um treinamento relativo mais rápido.

Figura 20 – Tempo computacional para treinamento e teste dos modelos



**Fonte:** Elaborada pelo autor (2022)

Considerando o tempo de teste, a relação entre SARIMA e CODYS inverte. Aquele foi mais rápido que este, o que é esperado devido a necessidade de múltiplas previsões do CODYS, mas mesmo com múltiplas execuções ele foi mais rápido que as redes neurais (MLP e LSTM).

## 7 CONCLUSÃO

O problema de previsão dos casos epidemiológicos, como a COVID-19, é multifatorial e complexo, o que leva a muitos erros e falhas nos modelos preditivos (UCHIDA; YOSHIDA, 2022). Um dos problemas para essas previsões é que se tratam de séries não estacionárias sujeitas a mudança de conceito.

A hipótese de pesquisa foi abordar o problema da previsão identificando a mudança de conceito nas séries de casos diários da pandemia e utilizar essa informação para treinar modelos especialistas para compor um sistema de múltiplos preditores e, por meio da seleção dinâmica, capaz de identificar conceitos no conjunto de teste e escolher os modelos candidatos mais adequados ao conceito vigente no momento da previsão.

O sistema proposto, CODYS, conseguiu resultados superiores a um conjunto de modelos de clássicos de séries temporais e aprendizado de máquina utilizados na literatura sobre o tema, tendo sido capaz de identificar regiões de competência com distribuições estatisticamente distintas, e os modelos especialistas treinados estavam mais aptos, em conjunto, a lidar com as mudanças de conceito das séries do que os modelos individuais consagrados na literatura.

A inovação deste trabalho é combinar a mudança de conceito com a seleção dinâmica para a construção de um sistema de previsão capaz de superar o desafio de prever séries com múltiplas variações de conceito, como as séries da COVID-19. Até o momento, essa abordagem não é encontrada na literatura sobre o tema conduzindo, portanto, a uma nova linha de pesquisa em previsão de séries temporais.

Algumas limitações da pesquisa foram: i) a previsão um passo a frente, ii) a limitação na quantidade de dados das séries, iii) modelos utilizados na composição do sistema proposto, e iv) os métodos de seleção dinâmica e combinação dos resultados.

O horizonte de previsão de um passo a frente implica um prazo de curto de previsão, portanto não permite afirmar que o sistema proposto apresenta desempenho superior em previsões de médio e longo prazo ou múltiplos passos a frente, nem quais adaptações seriam necessárias para tanto.

O tamanho das séries estudadas, 709 pontos de dados no total, podem ter prejudicado o desempenho modelos que requerem uma quantidade elevada de dados para um aprendizado eficaz e também limitado as opções de modelos para compor o sistema proposto.

O uso do modelo SARIMA como base para os modelos especialistas pode ter reduzido

o resultado atingido pelo sistema proposto, pois a utilização de outros modelos de base e a combinação entre poderia aumentar o poder preditivo do sistema proposto.

A utilização do MSE da previsão das janelas vizinhas como método de seleção vizinhas permitiu obter bons resultados nos cenários avaliados, entretanto outras métricas ou algoritmos para seleção dos modelos poderiam identificar melhor o conceito vigente, e selecionar melhores candidatos para a previsão.

Trabalhos futuros podem aprofundar a pesquisa e explorar oportunidades de melhoria no sistema proposto:

1. Seleção Dinâmica: novos métodos e parâmetros para selecionar modelos
  - a) Substituir a seleção pelo menor de previsão erro na região de competência por um classificador. Por exemplo, uma rede neural que recebe as previsões da região de competência e retorna as respectivas probabilidades para cada modelo em relação ao padrão de teste.
  - b) Substituir a seleção pelo menor de previsão erro na região de competência por mecanismos de atenção, como classificadores ou preditores.
  - c) Utilizar outras métricas de erro como *huber loss* ou média móvel exponencial.
  - d) Substituir a métrica de erro por métricas de distâncias como *Dynamic Time Warping*.
2. *Pools* heterogêneos: combinar diferentes modelos preditivos, por exemplo SARIMA com SVR ou ELM.
3. Detecção de conceitos: novos detectores de mudança de conceito e configurações de hiperparâmetros.
4. Domínio: utilizar o sistema proposto em outros domínios de aplicação com mudança de conceito, como a previsão de séries financeiras.

## REFERÊNCIAS

- AKIBA, T.; SANO, S.; YANASE, T.; OHTA, T.; KOYAMA, M. Optuna: A next-generation hyperparameter optimization framework. In: *Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. [S.l.: s.n.], 2019.
- ALA'RAJ, M.; MAJDALAWIEH, M.; NIZAMUDDIN, N. Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. *Infectious Disease Modelling*, Elsevier BV, v. 6, p. 98–111, 2021. Disponível em: <<https://doi.org/10.1016/j.idm.2020.11.007>>.
- ALLEN, D. W. Covid-19 Lockdown Cost/Benefits: A Critical Assessment of the Literature. *International Journal of the Economics of Business*, Routledge, v. 29, n. 1, p. 1–32, 2022. Disponível em: <<https://doi.org/10.1080/13571516.2021.1976051>>.
- ARDABILI, S. F.; MOSAVI, A.; GHAMISI, P.; FERDINAND, F.; VARKONYI-KOCZY, A. R.; REUTER, U.; RABCZUK, T.; ATKINSON, P. M. Covid-19 outbreak prediction with machine learning. *Algorithms*, MDPI, v. 13, n. 10, p. 249, 2020.
- ARUNKUMAR, K.; KALAGA, D. V.; KUMAR, C. M. S.; CHILKOOR, G.; KAWAJI, M.; BRENZA, T. M. Forecasting the dynamics of cumulative covid-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-regressive integrated moving average (arima) and seasonal auto-regressive integrated moving average (sarima). *Applied soft computing*, Elsevier, v. 103, p. 107161, 2021.
- BANK, W. *World Development Report 2022: Finance for an Equitable Recovery*. [S.l.]: The World Bank, 2022.
- BERNAL, J. L.; ANDREWS, N.; GOWER, C.; ROBERTSON, C.; STOWE, J.; TESSIER, E.; SIMMONS, R.; COTTRELL, S.; ROBERTS, R.; O'DOHERTY, M. et al. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ*, v. 373, p. n1088, may 2021. Disponível em: <<http://www.bmj.com/content/373/bmj.n1088.abstract>>.
- BIFET, A.; GAVALDA, R. Learning from time-changing data with adaptive windowing. In: SIAM. *Proceedings of the 2007 SIAM international conference on data mining*. 2007. p. 443–448. Disponível em: <<https://epubs.siam.org/doi/abs/10.1137/1.9781611972771.42>>.
- BOULMAIZ, T.; GUERMOUI, M.; BOUTAGHANE, H. Impact of training data size on the lstm performances for rainfall–runoff modeling. *Modeling Earth Systems and Environment*, Springer, v. 6, p. 2153–2164, 2020.
- BOX, G. E.; HUNTER, W. H.; HUNTER, S. et al. *Statistics for experimenters*. [S.l.]: John Wiley and sons New York, 1978. v. 664.
- BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. *Time series analysis: forecasting and control*. [S.l.]: John Wiley & Sons, 2015.
- CAVALCANTE, R. C.; MINKU, L. L.; OLIVEIRA, A. L. I. Fedd: Feature extraction for explicit concept drift detection in time series. In: *2016 International Joint Conference on Neural Networks (IJCNN)*. [S.l.: s.n.], 2016. p. 740–747.

- CHAURASIA, V.; PAL, S. Application of machine learning time series analysis for prediction covid-19 pandemic. *Research on Biomedical Engineering*, Springer, p. 1–13, 2020.
- CHRISTIE, A.; HENLEY, S. J.; MATTOCKS, L.; FERNANDO, R.; LANSKY, A.; AHMAD, F. B.; ADJEMIAN, J.; ANDERSON, R. N.; BINDER, A. M.; CAREY, K. et al. Decreases in covid-19 cases, emergency department visits, hospital admissions, and deaths among older adults following the introduction of covid-19 vaccine—united states, september 6, 2020–may 1, 2021. *Morbidity and Mortality Weekly Report*, Centers for Disease Control and Prevention, v. 70, n. 23, p. 858, 2021.
- COMITO, C.; PIZZUTI, C. Artificial intelligence for forecasting and diagnosing covid-19 pandemic: A focused review. *Artificial Intelligence in Medicine*, Elsevier, p. 102286, 2022.
- COWPERTWAIT, P. S.; METCALFE, A. V. *Introductory time series with R*. [S.l.]: Springer Science & Business Media, 2009.
- CRUZ, R. M. O.; SABOURIN, R.; CAVALCANTI, G. D. C. META-DES.Oracle: Meta-learning and feature selection for dynamic ensemble selection. *Information Fusion*, v. 38, p. 84–103, 2017. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S156625351730101X>>.
- CUCINOTTA, D.; VANELLI, M. Who declares covid-19 a pandemic. *Acta Bio Medica: Atenei Parmensis*, Mattioli 1885, v. 91, n. 1, p. 157, 2020.
- DAI, Q.; SONG, G.; YAO, C. Several Novel Dynamic Ensemble Selection Algorithms for Time Series Prediction. *Neural Processing Letters*, v. 50, n. 2, p. 1789–1829, 2019. ISSN 1573-773X. Disponível em: <<https://doi.org/10.1007/s11063-018-9957-7>>.
- DASGUPTA, S.; PAPADIMITRIOU, C. H.; VAZIRANI, U. V. *Algorithms*. [S.l.]: McGraw-Hill Higher Education New York, 2008.
- DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine learning research*, JMLR. org, v. 7, p. 1–30, 2006.
- DEVARAJ, J.; ELAVARASAN, R. M.; PUGAZHENDHI, R.; SHAFIULLAH, G.; GANESAN, S.; JEYSREE, A. K.; KHAN, I. A.; HOSSAIN, E. Forecasting of covid-19 cases using deep learning models: Is it reliable and practically significant? *Results in Physics*, Elsevier, v. 21, p. 103817, 2021.
- DONG, E.; RATCLIFF, J.; GOYEA, T. D.; KATZ, A.; LAU, R.; NG, T. K.; GARCIA, B.; BOLT, E.; PRATA, S.; ZHANG, D.; MURRAY, R. C.; BLAKE, M. R.; DU, H.; GANJKHANLOO, F.; AHMADI, F.; WILLIAMS, J.; CHOUDHURY, S.; GARDNER, L. M. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: data collection process, challenges faced, and lessons learned. *The Lancet Infectious Diseases*, 2022. ISSN 1473-3099. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1473309922004340>>.
- FANG, Z.-g.; YANG, S.-q.; LV, C.-x.; AN, S.-y.; WU, W. Application of a data-driven xgboost model for the prediction of covid-19 in the usa: a time-series study. *BMJ open*, British Medical Journal Publishing Group, v. 12, n. 7, p. e056685, 2022.
- FIRMINO, P. R. A.; SALES, J. P. D.; JÚNIOR, J. G.; SILVA, T. A. D. A non-central beta model to forecast and evaluate pandemics time series. *Chaos, Solitons & Fractals*, Elsevier, v. 140, p. 110211, 2020.

FLORENSA, D.; MATEO, J.; SPAIMOC, R.; MIRET, C.; GODOY, S.; SOLSONA, F.; GODOY, P. Severity of covid-19 cases in the months of predominance of the alpha and delta variants. *Scientific Reports*, Nature Publishing Group, v. 12, n. 1, p. 1–6, 2022.

GONÇALVES, P. M.; de Carvalho Santos, S. G. T.; BARROS, R. S. M.; VIEIRA, D. C. L. A comparative study on concept drift detectors. *Expert Systems with Applications*, v. 41, n. 18, p. 8144–8156, 2014. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417414004175>>.

HARROU, F.; ZEROUAL, A.; DAIRI, A.; SUN, Y. Deep learning methods for forecasting covid-19 time-series data: A comparative study. *Chaos, Solitons & Fractals*, v. 140, p. 110121, 2020. ISSN 0960-0779.

HERBY, J.; JONUNG, L.; HANKE, S. A literature review and meta-analysis of the effects of lockdowns on covid-19 mortality-ii. 2022.

HERNANDEZ-MATAMOROS, A.; FUJITA, H.; HAYASHI, T.; PEREZ-MEANA, H. Forecasting of covid19 per regions using arima models and polynomial functions. *Applied Soft Computing*, Elsevier, v. 96, p. 106610, 2020.

HOLMES, E. C.; GOLDSTEIN, S. A.; RASMUSSEN, A. L.; ROBERTSON, D. L.; CRITS-CHRISTOPH, A.; WERTHEIM, J. O.; ANTHONY, S. J.; BARCLAY, W. S.; BONI, M. F.; DOHERTY, P. C. et al. The origins of sars-cov-2: A critical review. *Cell*, Elsevier, v. 184, n. 19, p. 4848–4856, 2021.

HSIANG, S.; ALLEN, D.; ANNAN-PHAN, S.; BELL, K.; BOLLIGER, I.; CHONG, T.; DRUCKENMILLER, H.; HUANG, L. Y.; HULTGREN, A.; KRASOVICH, E.; LAU, P.; LEE, J.; ROLF, E.; TSENG, J.; WU, T. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, v. 584, n. 7820, p. 262–267, 2020. ISSN 1476-4687. Disponível em: <<https://doi.org/10.1038/s41586-020-2404-8>>.

HUANG, C.; WANG, Y.; LI, X.; REN, L.; ZHAO, J.; HU, Y.; ZHANG, L.; FAN, G.; XU, J.; GU, X. et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *The lancet*, Elsevier, v. 395, n. 10223, p. 497–506, 2020.

KERMACK, W. O.; MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, The Royal Society London, v. 115, n. 772, p. 700–721, 1927.

KOURENTZES, N.; BARROW, D. K.; CRONE, S. F. Neural network ensemble operators for time series forecasting. *Expert Systems with Applications*, v. 41, n. 9, p. 4235–4244, 2014. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417413009834>>.

LIMA, A.; BARBOSA, J.; NETO, P. M.; OLIVEIRA, A. Hybrid swarm enhanced classifier ensembles. In: *Anais do XVIII Encontro Nacional de Inteligência Artificial e Computacional*. Porto Alegre, RS, Brasil: SBC, 2021. p. 314–325. ISSN 2763-9061. Disponível em: <<https://sol.sbc.org.br/index.php/eniac/article/view/18263>>.

MAALIW, R. R.; BALLERA, M. A.; MABUNGA, Z. P.; MAHUSAY, A. T.; DEJELO, D. A.; SEÑO, M. P. An ensemble machine learning approach for time series forecasting of covid-19

cases. In: IEEE. *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. [S.l.], 2021. p. 0633–0640.

MEDSKER, L. R. *Hybrid intelligent systems*. [S.l.]: Springer Science & Business Media, 2012.

MEHTA, S. H.; CLIPMAN, S. J.; WESOLOWSKI, A.; SOLOMON, S. S. Holiday gatherings, mobility and SARS-CoV-2 transmission: results from 10 US states following Thanksgiving. *Scientific Reports*, v. 11, n. 1, p. 17328, 2021. ISSN 2045-2322. Disponível em: <<https://doi.org/10.1038/s41598-021-96779-6>>.

MINKU, L. L.; WHITE, A. P.; YAO, X. The impact of diversity on online ensemble learning in the presence of concept drift. *IEEE Transactions on Knowledge and Data Engineering*, v. 22, n. 5, p. 730–742, 2010.

MINKU, L. L.; YAO, X. Ddd: A new ensemble approach for dealing with concept drift. *IEEE Transactions on Knowledge and Data Engineering*, v. 24, n. 4, p. 619–633, 2012.

MORAIS, L. R. de A.; GOMES, G. S. da S. Forecasting daily covid-19 cases in the world with a hybrid ARIMA and neural network model. *Applied Soft Computing*, Elsevier BV, v. 126, p. 109315, set. 2022. Disponível em: <<https://doi.org/10.1016/j.asoc.2022.109315>>.

NADIM, S. S.; GHOSH, I.; CHATTOPADHYAY, J. Short-term predictions and prevention strategies for COVID-19: A model-based study. *Applied Mathematics and Computation*, Elsevier BV, v. 404, p. 126251, set. 2021. Disponível em: <<https://doi.org/10.1016/j.amc.2021.126251>>.

OLIVEIRA, J. F. de; SILVA, E. G.; NETO, P. S. de M. A hybrid system based on dynamic selection for time series forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, IEEE, 2021.

OPITZ, D.; MACLIN, R. Popular ensemble methods: An empirical study. *Journal of artificial intelligence research*, v. 11, p. 169–198, 1999.

PANDEY, G.; CHAUDHARY, P.; GUPTA, R.; PAL, S. Seir and regression model based covid-19 outbreak predictions in india. *arXiv preprint arXiv:2004.00958*, 2020.

PENG, L.; YANG, W.; ZHANG, D.; ZHUGE, C.; HONG, L. Epidemic analysis of covid-19 in china by dynamical modeling. *arXiv preprint arXiv:2002.06563*, 2020.

RASHED, E. A.; HIRATA, A. Infectivity upsurge by covid-19 viral variants in japan: Evidence from deep learning modeling. *International journal of environmental research and public health*, MDPI, v. 18, n. 15, p. 7799, 2021.

REN, S.-Y.; WANG, W.-B.; GAO, R.-D.; ZHOU, A.-M. Omicron variant (b. 1.1. 529) of sars-cov-2: Mutation, infectivity, transmission, and vaccine resistance. *World journal of clinical cases*, Baishideng Publishing Group Inc, v. 10, n. 1, p. 1, 2022.

RIBEIRO, M. H. D. M.; SILVA, R. G. da; MARIANI, V. C.; COELHO, L. dos S. Short-term forecasting covid-19 cumulative confirmed cases: Perspectives for brazil. *Chaos, Solitons & Fractals*, Elsevier, v. 135, p. 109853, 2020.

SAEZ, M.; TOBIAS, A.; VARGA, D.; BARCELÓ, M. A. Effectiveness of the measures to flatten the epidemic curve of COVID-19. The case of Spain. *Science of The Total Environment*, v. 727, p. 138761, 2020. ISSN 0048-9697. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0048969720322786>>.

SCHLIMMER, J. C.; GRANGER, R. H. Incremental learning from noisy data. *Machine learning*, Springer, v. 1, n. 3, p. 317–354, 1986.

SERGIO, A. T.; LIMA, T. P. F. de; LUDERMIR, T. B. Dynamic selection of forecast combiners. *Neurocomputing*, v. 218, p. 37–50, 2016. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231216309687>>.

SILVA, E. G.; CAVALCANTI, G. D.; OLIVEIRA, J. F. L. de; NETO, P. S. de M. On the evaluation of dynamic selection parameters for time series forecasting. In: IEEE. *2020 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2020. p. 1–7.

SILVA, E. G.; NETO, P. S. G. D. M.; CAVALCANTI, G. D. C. A dynamic predictor selection method based on recent temporal windows for time series forecasting. *IEEE Access*, v. 9, p. 108466–108479, 2021.

SMITH, T. G. et al. *pmdarima: ARIMA estimators for Python*. 2017–. [Online; accessed <today>]. Disponível em: <<http://www.alkaline-ml.com/pmdarima>>.

TALKHI, N.; FATEMI, N. A.; ATAEI, Z.; NOOGHABI, M. J. Modeling and forecasting number of confirmed and death caused covid-19 in iran: A comparison of time series forecasting methods. *Biomedical Signal Processing and Control*, Elsevier, v. 66, p. 102494, 2021.

UCHIDA, T.; YOSHIDA, K. Concept drift in japanese covid-19 infection data. *Procedia Computer Science*, Elsevier, v. 207, p. 380–387, 2022.

WEBB, G.; HYDE, R.; CAO, H.; NGUYEN, H. L.; PETITJEAN, F. Characterizing concept drift. *Data Mining and Knowledge Discovery*, Springer, v. 30, n. 4, p. 964–994, 2016.

WIDMER, G.; KUBAT, M. Learning in the presence of concept drift and hidden contexts. *Machine learning*, Springer, v. 23, n. 1, p. 69–101, 1996.

WOLPERT, D. H. The lack of a priori distinctions between learning algorithms. *Neural computation*, MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , v. 8, n. 7, p. 1341–1390, 1996.

ZEROUAL, A.; HARROU, F.; DAIRI, A.; SUN, Y. Deep learning methods for forecasting covid-19 time-series data: A comparative study. *Chaos, Solitons & Fractals*, Elsevier, v. 140, p. 110121, 2020.

ŽLIOBAITĖ, I. Learning under concept drift: an overview. *arXiv preprint arXiv:1010.4784*, 2010.