

# Identificação de Dias Depressivos a Partir de Classificação com Séries Temporais

Guilherme Augusto Rocha de Figueiredo

<sup>1</sup>Universidade Federal de Viçosa (UFV) – Viçosa, MG – Brasil

guilherme.figueiredo@ufv.br

**Abstract.** *This study aims to classify days as depressive or non-depressive based on motor activity data collected from individuals with and without a depression diagnosis. Initially, a tabular classification was performed to establish a baseline performance. Thus, this study utilizes time series classification techniques to capture information lost in data compression and improve classification performance.*

**Resumo.** *Este trabalho visa classificar dias como depressivos ou não depressivos com base em dados de atividade motora coletados de indivíduos com e sem diagnóstico de depressão. Inicialmente, uma classificação tabular foi realizada para gerar uma base de desempenho. Assim, este estudo utiliza técnicas de classificação de séries temporais com o objetivo de capturar informações perdidas na compressão dos dados e melhorar o desempenho da classificação.*

## 1. Introdução

A depressão é um transtorno mental prevalente e uma das principais causas de incapacidade no mundo, afetando milhões de pessoas de todas as idades e circunstâncias sociais. Segundo a Organização Mundial da Saúde (OMS), estima-se que mais de 280 milhões de pessoas no mundo sofram de depressão, uma condição que impacta negativamente o bem-estar psicológico e físico dos indivíduos. Além dos sintomas psicológicos, como tristeza persistente e perda de interesse, a depressão também se manifesta em alterações no sistema biológico interno do indivíduo, incluindo mudanças nos padrões de atividade motora.

Estudos recentes indicam que os padrões de atividade motora podem servir como indicadores objetivos para monitorar e prever episódios depressivos. Durante períodos de depressão, pacientes tendem a apresentar alterações motoras noturnas e diurnas que podem ser capturadas através de dispositivos de monitoramento contínuo, como actígrafos. A análise desses dados, especialmente em formato de séries temporais, permite que padrões sutis, muitas vezes imperceptíveis em análises tabulares, sejam revelados e associados a estados de humor.

Neste contexto, a utilização de dados de atividade motora para a classificação de episódios depressivos representa uma área promissora de pesquisa. Este trabalho faz uso do banco de dados Depresjon [Garcia-Ceja et al. 2018], uma base que contém informações sobre a atividade motora de indivíduos diagnosticados com diferentes tipos de depressão (unipolar e bipolar) e de indivíduos sem diagnóstico. Os dados foram coletados continuamente por meio de actígrafos colocados no pulso direito dos participantes, registrando a atividade motora minuto a minuto durante vários dias. A variabilidade nas

características dos participantes, como estado civil, sexo, idade e situação de trabalho, assim como a hospitalização de alguns deles, garante uma riqueza de dados para a análise de padrões relacionados à depressão.

O objetivo deste estudo é investigar se modelos de classificação baseados em séries temporais são capazes de distinguir de maneira eficaz dias depressivos de dias não depressivos e, mais especificamente, identificar diferentes subtipos de depressão.

## **2. Trabalhos Relacionados**

A classificação de séries temporais tem se mostrado uma abordagem eficiente em diversos contextos, principalmente quando há a necessidade de capturar padrões dinâmicos ao longo do tempo. No campo da saúde, isso é particularmente importante, pois muitos fenômenos biológicos e comportamentais, como a atividade motora de pacientes com depressão, apresentam variações temporais que são cruciais para uma análise mais precisa.

Diversos estudos apontam que o uso de séries temporais pode melhorar significativamente o desempenho em tarefas de classificação. Um exemplo disso é o trabalho de Bagnall et al. [Bagnall et al. 2017], que compara algoritmos de classificação de séries temporais, mostrando que técnicas como o ROCKET (RandOm Convolutional KErnel Transform) e o Time Series Forest (TSF) tendem a superar abordagens tradicionais de aprendizado supervisionado em cenários com dados temporais, devido à sua capacidade de capturar padrões temporais complexos de forma eficiente.

No contexto de detecção de depressão, a análise de dados de séries temporais tem ganhado espaço. A pesquisa de Saeb et al. [Saeb et al. 2016] utiliza dados de smartphones para monitorar padrões de mobilidade e sono de indivíduos, mostrando que é possível identificar estados depressivos através da análise temporal desses padrões. Outro estudo relevante é o de Jacobson et al. [Jacobson et al. 2020], que emprega séries temporais para prever episódios de depressão com base em dados de actigrafia, semelhantes aos utilizados neste trabalho, ressaltando o potencial dos métodos baseados em séries temporais para capturar as flutuações diárias da atividade motora.

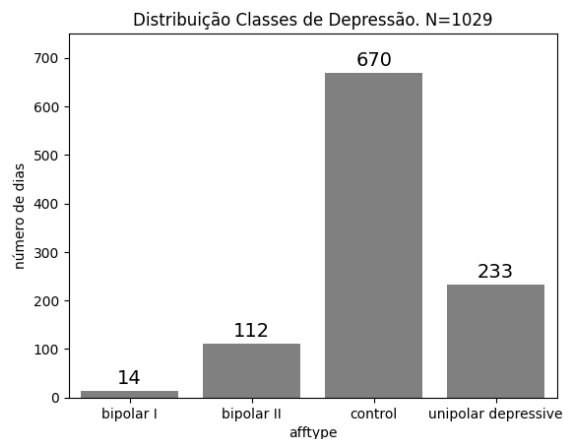
O dataset utilizado neste estudo foi inicialmente apresentado por Garcia-Ceja et al. [Garcia-Ceja et al. 2018], que investigaram como diferentes estados depressivos poderiam ser classificados a partir de dados de atividade motora coletados através de actígrafos. O estudo propôs uma abordagem de classificação tradicional utilizando dados tabulares e forneceu uma base de comparação para pesquisas subsequentes, incluindo esta, que faz uso de algoritmos para séries temporais a fim de explorar se essas técnicas podem melhorar o desempenho da classificação.

## **3. Metodologia**

### **3.1. Conjunto de Dados**

Os dados utilizados neste estudo foram coletados de indivíduos com e sem diagnóstico de depressão, registrando sua atividade motora ao longo de vários dias. O dataset inclui quatro grupos distintos: depressão unipolar, bipolar I, bipolar II e um grupo controle de indivíduos sem depressão. Cada série temporal é composta pela atividade motora do indivíduo registrada minuto a minuto, permitindo uma análise detalhada dos padrões diários de movimento.

Cada indivíduo teve sua série temporal dividida em segmentos diários, resultando em diferentes séries temporais menores para cada dia, com dias incompletos sendo descartados. O dataset resultante contém 1029 séries independentes de tamanho 1440, distribuídas de forma não uniforme entre os grupos de análise. A figura 1 mostra tal distribuição.



**Figura 1. Distribuição das séries temporais entre as diferentes classes de análise.**

O conjunto de dados foi utilizado em duas classificações distintas. A primeira consiste em uma classificação binária entre dias depressivos e não depressivos, enquanto a segunda remove o grupo de controle e busca identificar o tipo específico de depressão daquele dia.

### 3.2. Algoritmos de Classificação

Neste trabalho, foram exploradas duas técnicas de classificação de séries temporais: ROCKET (RandOm Convolutional KErnel Transform) e Time Series Forest. A escolha desses algoritmos foi motivada por sua eficácia comprovada em tarefas de classificação de séries temporais e por sua eficiência em termos de tempo de processamento e desempenho.

O ROCKET é uma técnica que utiliza convoluções aleatórias para extrair características relevantes de séries temporais. Sua simplicidade e rapidez na extração de características o tornam particularmente adequado para o problema em questão, onde é preciso detectar padrões em dados de alta frequência, capturados minuto a minuto, e com grandes variações.

O Time Series Forest (TSF) é um algoritmo de aprendizado por ensemble que combina diversas árvores de decisão para modelar padrões temporais de maneira robusta. Essa abordagem é particularmente eficaz em identificar variações ao longo do tempo, conseguindo lidar tanto com a complexidade dos dados quanto com flutuações sutis nos padrões de atividade motora. Além disso, o TSF se destaca por sua capacidade de capturar informações relevantes em diferentes janelas de tempo, o que o torna adequado para dados temporais heterogêneos e de difícil modelagem.

### 3.3. Validação e Métricas de Avaliação

Para validar a eficácia dos modelos, foi aplicada uma validação cruzada 10-Fold, assim como na análise tabular original do dataset. Esse método de validação garante que to-

dos os dados sejam utilizados tanto para treino quanto para teste, reduzindo o risco de overfitting e proporcionando uma avaliação robusta do desempenho dos modelos.

Além disso, as mesmas métricas utilizadas na análise original foram aplicadas para avaliar os modelos neste estudo: Precisão (Precision), Revocação (Recall), Acurácia (Accuracy), Especificidade (Specificity), Coeficiente de Correlação de Matthews (Matthews correlation coefficient) e F1-score. Essas métricas fornecem uma visão abrangente da performance dos modelos, considerando tanto a capacidade de identificar corretamente os dias depressivos quanto a habilidade de evitar falsos positivos e falsos negativos.

### 3.4. Implementação e Ajuste de Hiperparâmetros

O estudo foi desenvolvido em Python, utilizando os algoritmos implementados pela biblioteca `sktime`. Outras partes do processo de classificação, como a normalização dos dados e as métricas de avaliação, foram realizadas utilizando a biblioteca `sklearn`. As séries temporais passaram por uma normalização Min-Max antes da classificação, ajustando os valores de atividade motora para o intervalo entre 0 e 1.

Além disso, o método `GridSearch` foi utilizado para a escolha dos hiperparâmetros. No caso do `ROCKET`, o aumento do número de kernels não resultou em uma melhora significativa no desempenho, portanto, o valor padrão de 10.000 kernels foi mantido. A melhor configuração para o `Time Series Forest` foi obtida com os hiperparâmetros `n_estimators=500` e `min_interval=45`.

## 4. Resultados

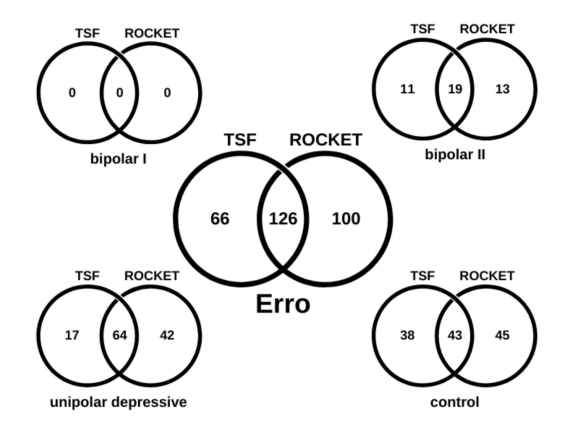
### 4.1. Classificação Binária

**Tabela 1. Resultados da classificação entre dias depressivos e não depressivos.**

		Precision	Recall	Accuracy	MCC	F1-Score	Specificity
Classifier	Subclass						
Time Series Forest	depressed	0.756571	0.690873	0.813383	0.583761	0.721203	0.879104
	nondepressed	0.841528	0.879104	0.813383	0.583761	0.859617	0.690873
	weighted average	<b>0.811943</b>	<b>0.813383</b>	<b>0.813383</b>	<b>0.583761</b>	<b>0.811328</b>	<b>0.756595</b>
ROCKET	depressed	0.714866	0.615556	0.780354	0.503494	0.660512	0.868657
	nondepressed	0.808948	0.868657	0.780354	0.503494	0.837518	0.615556
	weighted average	0.776145	0.780354	0.780354	0.503494	0.775773	0.703858

A Tabela 1 apresenta os resultados da classificação binária entre dias depressivos e não depressivos, enquanto a Figura 2 ilustra a distribuição dos erros de predição dos algoritmos entre as diferentes classes de depressão.

Na classificação binária, ambos os modelos para séries temporais superaram todos os algoritmos que utilizam dados tabulares, conforme apresentado no artigo original,

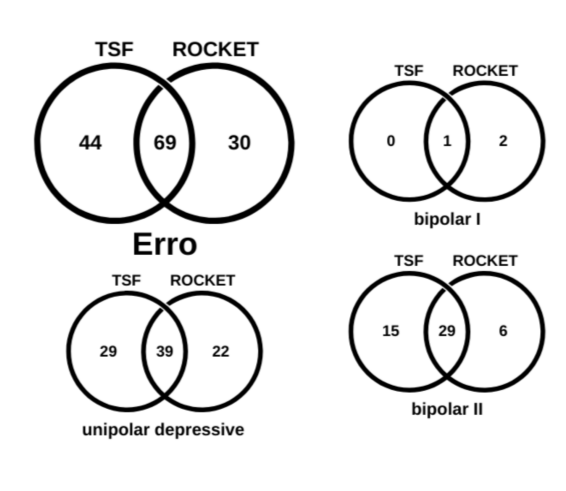


**Figura 2.** Representação, em diagrama de venn, da distribuição do erro dos algoritmos, na classificação entre dias depressivos e não depressivos, entre os diferentes tipos de depressão.

exceto pela métrica de especificidade no ROCKET. Observa-se, ao analisar os valores das classes para essa métrica, que o ROCKET tende a gerar mais falsos positivos na classe de dias não depressivos. O Time Series Forest apresentou os melhores resultados em quase todas as métricas, exceto na Revocação para a classe de dias depressivos, onde o desempenho foi inferior a mais da metade dos algoritmos tabulares. Isso indica que, apesar do uso de séries temporais e melhora na classificação de dias não depressivos, os algoritmos ainda enfrentam dificuldades para capturar os dias depressivos de forma eficaz.

Em relação aos diferentes tipos de depressão, os diagramas de Venn revelam que ambos os algoritmos acertaram todas as instâncias da classe bipolar I, apesar de esta ser a menor classe do conjunto de dados. Por outro lado, o desempenho na identificação da classe unipolar depressivo foi insatisfatório, com mais de 50% das amostras sendo classificadas de forma errônea.

#### 4.2. Classificação dos Tipos de Depressão



**Figura 3.** Representação, em diagrama de venn, da distribuição do erro dos algoritmos na classificação entre os diferentes tipos de depressão.

**Tabela 2. Resultados da classificação entre os diferentes tipos de depressão.**

		Precision	Recall	Accuracy	MCC	F1-Score	Specificity
Classifier	Subclass						
Time Series Forest	bipolar I	0.516667	0.550000	0.975000	0.519826	0.513333	0.994202
	bipolar II	0.551667	0.214394	0.693492	0.182465	0.302161	0.911000
	unipolar depressive	0.707370	0.922464	0.701825	0.288552	0.800440	0.293590
	weighted average	0.650728	0.685159	0.685159	0.267916	0.632357	0.513633
ROCKET	bipolar I	0.300000	0.250000	0.972143	0.269693	0.266667	1.000000
	bipolar II	0.613276	0.374242	0.726905	0.309896	0.458299	0.886500
	unipolar depressive	0.751642	0.918841	0.749206	0.422319	0.826001	0.434615
	weighted average	<b>0.693202</b>	<b>0.724127</b>	<b>0.724127</b>	<b>0.370611</b>	<b>0.691086</b>	<b>0.596988</b>

A Tabela 2 e a Figura 3 apresentam os resultados da classificação entre os diferentes tipos de depressão no dataset, utilizando o mesmo formato da classificação binária.

Ao contrário dos resultados obtidos na classificação binária, ao tentar separar as classes de depressão, o ROCKET apresentou o melhor desempenho em todas as métricas, errando mais amostras apenas da classe bipolar I, menor classe do dataset. Isso evidencia a capacidade de seus kernels na detecção de nunces e extração de características das time series, capturando as diferentes variações entre os tipos de depressão.

O Time Series Forest obteve melhores resultados na classificação das instâncias do tipo bipolar I, porém, para alcançar tal feito, houve um aumento nos erros de predição das outras classes, especialmente da classe bipolar II, que obteve valores insatisfatórios de Revocação e MCC. Esse desbalanceamento ocorre especialmente devido a distribuição não uniforme entre as classes no dataset. Como a classe de controle foi removida nesta classificação, o tipo unipolar depressive se tornou a classe predominante, refletindo em uma melhora na sua identificação pelos algoritmos.

## 5. Conclusão

A classificação utilizando séries temporais demonstrou ser superior à abordagem tabular, gerando melhores resultados na maioria dos cenários. No entanto, os algoritmos ainda apresentam desafios ao capturar adequadamente as amostras de dias depressivos. Estudos futuros poderiam se concentrar em entender melhor os fatores que levam à classificação de um dia como depressivo, seja por meio de técnicas de explicabilidade dos modelos ou pelo uso de *shaplets* para descrever com maior precisão esses dias.

Adicionalmente, os algoritmos mostraram desempenho superior em diferentes cenários: Time Series Forest na classificação binária e ROCKET na classificação multiclasse. Logo, podemos concluir que o TSF apresentou uma melhor generalização para identificar um dia depressivo, enquanto o ROCKET conseguiu diferenciar de forma mais precisa os diferentes tipos de depressão. Isso sugere que uma combinação hierárquica

desses métodos pode ser uma abordagem promissora para melhorar ainda mais os resultados.

## Referências

- Bagnall, A., Lines, J., Bostrom, A., Large, J., and Keogh, E. (2017). The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, 31(3):606–660.
- Garcia-Ceja, E., Riegler, M., Jakobsen, P., Tørresen, J., Nordgreen, T., Oedegaard, K. J., and Fasmer, O. B. (2018). Depresjon: A motor activity database of depression episodes in unipolar and bipolar patients. In *Proceedings of the 9th ACM on Multimedia Systems Conference (MMSys'18)*, pages 472–477, Amsterdam, The Netherlands. ACM.
- Jacobson, N. C., Weingarden, H., and Wilhelm, S. (2020). Predicting moment-to-moment physical activity and mental health: An idiographic mobile sensing approach. *Internet Interventions*, 20:100314.
- Saeb, S., Zhang, M., Kording, K. P., and Mohr, D. C. (2016). The relationship between mobile phone location sensor data and depressive symptom severity. *JMIR mHealth and uHealth*, 3(1):e2316.