

Análise de Algoritmos de Aprendizado de Máquina Aplicados à Qualidade do Sono

Guilherme Eduardo Gonçalves da Silva
Universidade Federal do Paraná
Curitiba, Paraná, BR
guilherme.eduardo1@ufpr.br

Henrique Margotte
Universidade Federal do Paraná
Curitiba, Paraná, BR
hmargotte@inf.ufpr.br

Heloísa Dias Viotto
Universidade Federal do Paraná
Curitiba, Paraná, BR
heloisa.viotto@ufpr.br

Aurora Trinidad Ramirez Pozo
Universidade Federal do Paraná
Curitiba, Paraná, BR
aurora@inf.ufpr.br

ABSTRACT

Sleep plays an important role in people's quality of life, and its deprivation can lead to disorders such as insomnia and sleep apnea. In this study, we evaluate a dataset related to sleep quality from 374 users and use it to compare classification models with hyperparameter tuning. The cross-validation technique was employed in order to obtain the best accuracy, F1-score, precision and recall metrics. The results show that the Random Forest model is the most promising for this dataset, achieving an accuracy and F1-score values of approximately 92% and 90%, respectively.

KEYWORDS

Sono; Qualidade do Sono; Modelos de Classificação; Aprendizado de Máquina

1 INTRODUÇÃO

O sono é essencial para a saúde, sendo o momento que o corpo encontra para realizar manutenções, além de promover o seu desenvolvimento, sendo de extrema importância para o bem-estar físico e mental. O comprometimento desse estado natural pode afetar diversos mecanismos do corpo, como o sistema neural, imunológico, cardiovascular e psicológico.

Existem dois tipos de fases que constituem parte do sono, sendo elas: o movimento rápido dos olhos (REM) e o sono de movimento não-rápido (NREM), a etapa mais ativa. Nosso corpo passa por ciclos de 90 a 120 minutos de 4 a 6 vezes por noite e a cada transição o corpo reage a mudanças fisiológicas a fim de trazer benefícios para o corpo [1]. Os dados quantitativos e qualitativos do sono podem ser obtidos com o exame de polissonografia.

A qualidade do sono pode ser alterada por diversos fatores, como envelhecimento, depressão, uso de medicamentos, entre outros [2], podendo resultar em um aumento na gravidade dos distúrbios do sono, que, por sua vez, estão frequentemente associados ao abuso de substâncias, doenças crônicas, condições psiquiátricas, sedentarismo e ao declínio da memória com o avanço da idade [3].

O objetivo desse estudo é propor uma análise em uma base de dados denominada Saúde do Sono e Estilo de vida (*Sleep Health and Lifestyle dataset*) [4], o qual possui informações sobre 374 usuários sintéticos, incluindo dados como idade, gênero e ocupação, além de informações relacionadas ao sono, como duração, qualidade e distúrbio.

O experimento tem como finalidade avaliar diferentes modelos de classificação com o intuito de selecionar aqueles que apresentam melhores resultados de acurácia, *f1-score*, precisão e sensibilidade. O método de avaliação aplicado é a técnica de validação cruzada (*Cross Validation*) juntamente com a aplicação de hiperparâmetros específicos em cada modelo para auxílio no aprendizado. O código está disponível em <https://github.com/Guilherme-Eduardo/Trab-Machine-Learning>.

2 TRABALHOS CORRELATOS

Na Tabela 1, é possível visualizar os resultados de outros pesquisadores obtidos por meio da mesma base de dados. A técnica *stacking* (ST) foi uma abordagem introduzida apenas neste estudo, sendo assim, não foi incluída na tabela.

Tabela 1: Acurácias (%) obtidas por outros trabalhos.

Artigos	Modelos				
	NB	KNN	DT	RF	SVM
Alshumayri et. al [5]	-	86,67	-	92,15	90,02
Taher et. al [6]	-	-	-	0,9026	-
Migliaccio et. al [7]	-	-	85,00	92,50	87,50
Prasad et. al [8]	-	97,00	96,00	-	90,00
Alshammari et. al [9]	-	83,19	88,50	91,15	92,04
Saini et. al [10]	-	-	-	90,27	90,99
Kuamr et. al [11]	53,00	88,00	-	90,00	-

É possível observar que Prasad et. al [8] conseguiu resultados significativos para os modelos *k-Nearest Neighbors* (KNN) e *Árvore de decisão* (DT). Já Migliaccio et. al [7] e Alshammari et. al [9] obtiveram as maiores acurácias para os modelos *Random Forest* (RF) e *Support Vector Machine* (SVM), respectivamente. Por fim, Kuamr et. al [11] foi o único trabalho que testou essa mesma base de dados no *Naive Bayes* (NB).

3 SOLUÇÃO PROPOSTA

A solução proposta aplica os modelos clássicos de Aprendizado de Máquina para a classificação de distúrbios relacionados ao sono.

3.1 Conjunto de Dados

O conjunto de dados sintéticos utilizados possui características relacionado ao sono e hábitos diários de 374 pessoas. Os dados presentes são:

- Gênero;
- Idade;
- Ocupação;
- Duração de Sono;
- Qualidade do Sono;
- Nível de Atividade Física;
- Nível de Estresse;
- Índice de Massa Corporal (BMI);
- Pressão Sanguínea;
- Batimentos Cardíacos;
- Passos Diários; e,
- Distúrbio de Sono.

Há três classificações de distúrbios de sono: normal, insônia e apneia. O normal é o sono saudável. A insônia está relacionada com a dificuldade de dormir e manter-se neste estado, tendo o diagnóstico relacionado à dificuldade de foco, retenção de memória e entre outros. [12]. Por fim, a apneia, também conhecida como apneia obstrutiva do sono, é causado pela obstrução completa das vias aéreas superiores, ocasionando despertares e sono não reparador. [13]

3.2 Pré-Processamento

O pré-processamento foi realizado com o objetivo de preparar o conjunto de dados para o uso dos modelos de aprendizado. O atributo denominado de pressão sanguínea foi dividida em duas colunas, sendo representado por valores de pressão mínima e máxima. Além disso, a coluna de distúrbio de sono contendo as informações como Normal, Insônia e Apneia do Sono foram convertidas em valores numéricos. Por fim, os atributos nominais, como Ocupação e Gênero e BMI foram convertidos em valores binários por meio de *One-Hot Encoding*.

3.3 Modelos

Este trabalho utiliza 7 modelos, sendo eles:

- NB: classificador baseado no teorema de Bayes que utiliza a probabilidade condicional para encontrar uma probabilidade a posteriori dada uma a priori. Há alguns tipos, sendo o Gaussiano o utilizado nesse trabalho. [14, 15]
- KNN: modelo de classificação e regressão que mapeia os dados de treino em um plano e, ao obter uma amostra desconhecida, classifica-a de acordo com os K valores mais próximos dela. [16]
- DT: utiliza uma árvore construída a partir dos dados para formar uma série de questões que são desdobradas em nós binários ou múltiplos, utilizando o princípio de dividir e conquistar. [17]
- SVM: procura um hiperplano que divida os dados mantendo a maior margem possível entre os pontos mais próximos desta divisão utilizando uma função de kernel. [18]
- RF: método mais complexo de aprendizado de máquina que utiliza mais de uma árvore de decisão para tomar um veredito. [19]
- *Multilayer Perceptron* (MLP): rede neural formada por três tipos de camadas: entrada, a qual recebe os dados, oculta, a qual processa os dados, e saída, a qual encontra uma classe [16]. Neste estudo foi utilizado uma MLP simples de duas

camadas ocultas compostas por 32 e 16 neurônios, ambas com função de ativação ReLU.

- ST: combina a saída de múltiplos classificadores mais simples, utilizando o aprendizado em camadas. [20]

3.4 Hiperparâmetros

Para obter melhores resultados, foi definido hiperparâmetros em cada modelo, sendo configurações externas e que são usadas antes do treinamento para auxiliar no aprendizado. Para isso, foi definido valores arbitrários para cada parâmetro. Abaixo apresenta a melhor combinação de hiperparâmetros para cada modelo.

- KNN: $n_neighbors=7$;
- NB: padrão da biblioteca *scikit-learn*;
- DT: $criterion=gini$;
- RF: $max_depth=150$ com $criterion=Entropy$.
- SVM: $kernel=rbf$ com $gamma=0.01$ e $C=10$.
- MLP: Otimizador = *Adam*, $criterion=Cross\ Entropy\ Loss$ e uma taxa de aprendizagem de 10^{-3} .

Para o ST, foram utilizados quatro modelos de base: KNN, DT, NB e o SVM, sendo um deles empregado como *final_estimator*. Todos os modelos utilizados neste método foram utilizados com a configuração de hiperparâmetros padrão da biblioteca *scikit-learn*.

3.5 Métricas

Como métricas, escolhemos utilizar a acurácia, precisão, sensibilidade e o *F1-score*, apresentadas nas Equações 1, 2 e 3, para definir o melhor modelo por serem utilizadas nos trabalhos observados. Os resultados apresentados correspondem à média aritmética de 5 execuções, seguindo o método de validação cruzada. Abaixo as equações são descritas, em que *VP*, *FP* e *FN* representam, respectivamente, verdadeiros positivos, falsos positivos e falsos negativos.

$$\text{Acurácia} = \frac{\text{Total de Acertos}}{\text{Total de Itens}} \quad (1)$$

$$\text{Precisão} = \frac{VP}{VP + FP} \quad \text{Sensibilidade} = \frac{VP}{VP + FN} \quad (2)$$

$$F1\text{-score} = \frac{2 \cdot \text{Precisão} \cdot \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}} \quad (3)$$

4 RESULTADOS ALCANÇADOS

Na Tabela 2, são apresentados a média e o desvio-padrão das métricas mencionadas anteriormente para os modelos utilizados neste trabalho. De forma geral, todos os modelos alcançaram resultados satisfatórios em todas as métricas, com destaque para o RF, sendo este o que alcançou os melhores valores.

Seguido do RF, encontra-se o SVM e o MLP, com acurácia de 87%. O SVM destaca-se pelos bons valores de sensibilidade e precisão, mostrando que possui uma boa capacidade discriminativa, conseguindo reconhecer instâncias positivas e evitar predições equivocadas. Já o MLP apresenta desempenho equilibrado, no entanto, também foi o modelo que apresentou o maior desvio padrão, ou seja, é um modelo com maior variabilidade nos resultados.

Uma novidade deste trabalho em relação a outros analisados é o uso da técnica *Stacking*. No entanto, ela não trouxe bons ganhos, não

Tabela 2: Média e Desvio Padrão (%) dos Resultados Alcançados pelos Modelos.

Modelo	Acurácia	Precisão	Sensibilidade	F1-Score
KNN	89,30 (4,82)	86,78 (3,57)	86,67 (3,75)	86,50 (3,50)
Naive Bayes	89,04 (3,51)	85,89 (5,31)	86,89 (5,97)	86,19 (5,55)
Árvore de Decisão	90,11 (2,46)	90,11 (2,46)	86,81 (3,62)	87,20 (3,12)
SVM	91,72 (4,24)	90,11 (5,58)	89,13 (6,38)	87,45 (3,80)
MLP	90,37 (4,93)	87,86 (5,19)	88,06 (6,74)	87,64 (5,81)
Random forest	91,98 (3,03)	90,32 (3,44)	89,54 (4,38)	89,84 (3,91)
Stacking (KNN)	88,24 (4,41)	89,34 (3,98)	86,60 (5,49)	87,49 (4,70)
Stacking (SVM)	90,11 (2,61)	89,04 (3,86)	86,60 (4,22)	87,45 (3,80)
Stacking (NB)	89,30 (3,17)	86,03 (3,63)	86,78 (4,63)	86,32 (4,30)
Stacking (DT)	81,57 (9,35)	78,11 (9,81)	78,11 (9,81)	78,27 (9,69)

atingindo o melhor modelo e, em alguns casos, piorando os resultados alcançados pelos modelos individualmente. O ST utilizando o NB e o KNN como *final_estimator* trouxe piora nos resultados alcançados pelos modelos individualmente, além do ensemble utilizando a árvore de decisão apresentou o menor desempenho.

Por fim, também destaca-se o modelo NB, o qual apresentou um resultado maior que o encontrado na literatura.

5 CONSIDERAÇÕES FINAIS

Com este trabalho foi possível realizar uma comparação com os diversos modelos de aprendizado de máquina, os quais foram aplicados sobre um conjunto de dados sobre a qualidade do sono em diferentes estilos de vida. Sendo assim, o modelo RF e SVM apresentaram os melhores desempenhos, alcançando uma acurácia aproximada de 92%, destacando-se como a abordagem mais eficiente para o problema proposto. Além disso, apresentamos métricas para o ST e um resultado melhor para o NB em relação aos outros trabalhos científicos analisados.

Como trabalhos futuros, pode-se citar o uso de modelos mais complexos, como uma versão maior do MLP e o XGBoost, além de outros métodos de ensemble como o *blending*. Ajustes no pré-processamento também podem ser realizados, além de aumentar o conjunto de dados através da síntese de dados.

REFERÊNCIAS

- [1] Arpana Prasad, V Asha, M T Vasumathi, A. Kalaivani, Aniket Kumar, and Arpita S Kulkarni. Machine learning approach to sleep stage classification based on eeg signal. In *2025 International Conference on Visual Analytics and Data Visualization (ICVADV)*, pages 1–6, 2025. doi: 10.1109/ICVADV63329.2025.10961707.
- [2] AK Patel, V Reddy, KR Shumway, et al. Physiology, sleep stages. In *StatPearls [Internet]*. StatPearls Publishing, 2025. URL https://www.ncbi.nlm.nih.gov/translate.google/books/NBK526132/?x_tr_sl=en&x_tr_tl=pt&x_tr_hl=pt&x_tr_pto=tc. [Updated 2024 Jan 26].
- [3] EE Jaqua, M Hanna, W Labib, C Moore, and V Matossian. Common sleep disorders affecting older adults. *Perm J*, 27(1):122–132, 2023. doi: 10.7812/TPP/22.114. URL <https://doi.org/10.7812/TPP/22.114>. Epub 2022 Dec 12.
- [4] Laksika Tharmalingam. Sleep health and lifestyle dataset. Kaggle, 2023. URL <https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset>. Accessed: 2025-12-12.
- [5] Sarah Alshumayri, Reema Abdallah, Yehya Asseri, and Zain Balfagih. Sleep disorder analysis: Unveiling the interplay between lifestyle health and sleep quality. In *2024 21st Learning and Technology Conference (LT)*, pages 149–154, 2024. doi: 10.1109/LT60077.2024.10468720.
- [6] Abrar Taher and Walid Ibn Zinnah Ayon. Exploring sleep disorders: A comparative analysis of machine learning algorithms on sleep health and lifestyle data. In *2024 IEEE International Conference on Power, Electrical, Electronics and Industrial Applications (PEELACON)*, pages 71–75, 2024. doi: 10.1109/PEELACON63629.2024.10800593.
- [7] Maddalena Migliaccio, Andrea Abate, Carmen Bisogni, and Lucia Cimmino. Machine learning-driven stress prediction: A comparative analysis and web application using the sleep health and lifestyle dataset. In *2025 6th International Conference on Bio-engineering for Smart Technologies (BioSMART)*, pages 1–4, 2025. doi: 10.1109/BioSMART66413.2025.11046172.
- [8] Arpana Prasad, V Asha, M T Vasumathi, A. Kalaivani, Aniket Kumar, and Arpita S Kulkarni. Machine learning approach to sleep stage classification based on eeg signal. In *2025 International Conference on Visual Analytics and Data Visualization (ICVADV)*, pages 1–6, 2025. doi: 10.1109/ICVADV63329.2025.10961707.
- [9] Talal Sarheed Alshammari. Applying machine learning algorithms for the classification of sleep disorders. *IEEE Access*, 12:36110–36121, 2024. doi: 10.1109/ACCESS.2024.3374408.
- [10] Sandeep Kumar Saini, Garima Chandel, Aldrin Shayamano, and Ologe Temitope Jemima. Identification of sleep disorder using machine learning. In *2024 International Conference on Futuristic Technologies in Control Systems Renewable Energy (ICFCR)*, pages 1–6, 2024. doi: 10.1109/ICFCR64128.2024.10762998.
- [11] Mukesh Kumar, Bilal Ahmed, Hari Mohan Mishra, Amit Kumar Jha, Pushendra Kumar Sikarwal, and Sahil Rampal. Predictive sleep disorder modelling: Using machine learning with lifestyle and sleep health data. In *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, pages 1–7, 2024. doi: 10.1109/ACCAI61061.2024.10602153.
- [12] Diksha Bhartia, Lavanya G Prabhu, Darshan Nakshatri, Kaliraj S, and G. Pradeep Reddy. Insomnia detection using machine learning. In *2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, pages 915–919, 2025. doi: 10.1109/ICICV64824.2025.11085974.
- [13] John M. Slowik, Abdulghani Sankari, and Jeffrey F. Collen. *Obstructive Sleep Apnea*. StatPearls Publishing, Treasure Island (FL), March 2025. URL <https://www.ncbi.nlm.nih.gov/books/NBK459252/>. [Updated 2025 Mar 4].
- [14] Manisa Manoswini, Biswajit Sahoo, and Aleena Swetapadma. Language impairment detection based on speech signals and gaussian naive bayes algorithm. In *2024 OITS International Conference on Information Technology (OCIT)*, pages 310–314, 2024. doi: 10.1109/OCIT65031.2024.00061.
- [15] Vijay Vijay and Pushpneel Verma. Variants of naive bayes algorithm for hate speech detection in text documents. In *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*, pages 18–21, 2023. doi: 10.1109/AISC56616.2023.10085511.
- [16] Sneha Potghan, R. Rajamenakshi, and Archana Bhise. Multi-layer perceptron based lung tumor classification. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pages 499–502, 2018. doi: 10.1109/ICECA.2018.8474864.
- [17] Shovan Chowdhury and Marco P. Schoen. Research paper classification using supervised machine learning techniques. In *2020 Intermountain Engineering, Technology and Computing (IETC)*, pages 1–6, 2020. doi: 10.1109/IETC47856.2020.9249211.
- [18] Shireen Al-Ofary and Hamza Osman Ilhan. Classification of pca based reduced deep features by svm for diagnosing lung and colon cancer. In *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, pages 1–5, 2023. doi: 10.1109/HORA58378.2023.10156720.
- [19] Yixiao Gu. A comparative analysis study of stock prediction based on random forest and decision tree. In *2024 International Conference on Electronics and Devices, Computational Science (ICEDCS)*, pages 96–100, 2024. doi: 10.1109/ICEDCS64328.2024.00022.
- [20] Nidhi Ramteke and Priya Maidamwar. Cardiac patient data classification using ensemble machine learning technique. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–6, 2023. doi: 10.1109/ICCCNT56998.2023.10307702.