Uso Problemático de Mídias Sociais e Falta de Atenção: Uma Análise via Arquitetura de Processamento de Linguagem Natural Híbrida

Isabelle Santiago Cardoso Rizzo[1], **Silas Lima Filho**[1]

[1]Programa de Pós-Graduação em Informática
Universidade Federal do Rio de Janeiro
Rio de Janeiro, RJ – Brasil

irizzo@ufrj.br, silaslfilho@ic.ufrj.br

Abstract. *Problematic social media use severely compromises executive functions and attention. To investigate this phenomenon, an analytical framework is proposed to identify behavioral patterns. As a proof of concept, a Natural Language Processing pipeline was developed using YouTube data. Its methodological innovation lies in a hybrid architecture (Sentence Transformers and Large Language Models) capable of classifying noisy text with high precision. Preliminary trials using topic modeling validated the approach by extracting clusters related to focus tactics and neurodivergence. We conclude that this methodology ensures large-scale data provenance and curation, yielding reliable datasets for digital health research.*

Resumo. *O uso problemático de mídias sociais compromete severamente funções executivas e a atenção. Para investigar esse fenômeno, propõe-se um framework analítico de identificação de padrões comportamentais. Como prova de conceito, desenvolveu-se um pipeline de Processamento de Linguagem Natural com dados do YouTube, cuja inovação reside numa arquitetura híbrida (Sentence Transformers e Large Language Models) capaz de classificar textos ruidosos com alta precisão. Ensaios preliminares via modelagem de tópicos validaram a abordagem ao extrair clusters sobre táticas de foco e neurodivergência. Conclui-se que a metodologia garante proveniência e curadoria de dados em larga escala, viabilizando datasets confiáveis para pesquisas em saúde digital.*

Informações Gerais. Nível: Mestrado. **Aluno:** Isabelle Santiago Cardoso Rizzo. **Orientador:** Silas Lima Filho. **Programa:** Programa de Pós-Graduação em Informática da Universidade Federal do Rio de Janeiro. **Admissão:** Março 2025. **Exame de qualificação:** Julho 2026. **Defesa prevista:** Fevereiro 2027. **Etapas concluídas:** Créditos obrigatórios em disciplinas, revisão de literatura, definição do problema e proposta.

1. Introdução

A ubiquidade das tecnologias insere a sociedade em uma dinâmica de hiperconexão, evidenciada no Brasil pelo uso diário da internet por 95,2% da população, majoritariamente via *smartphones* (98,8%), dispositivo possuído para uso pessoal por 88,9% da população com mais de 10 anos [IBGE 2025]. Nesse cenário, consolida-se a **economia da atenção**, em que a atenção e o tempo do usuário são disputados como recursos escassos por plataformas que combinam *Big Data*, Inteligência Artificial (IA), gamificação e algoritmos adaptativos para maximizar sua retenção [De La Torre et al. 2025, Desai and Thombre 2025].

Apesar de viabilizar inovações, essa infraestrutura digital pode induzir a comportamentos nocivos, como o Uso Problemático de Mídias Sociais (UPMS) e a Adicção em Mídias Sociais (AMS). Frequentemente tratados como sinônimos, descrevem um padrão persistente de baixo autocontrole e uso compulsivo e excessivo [Amirthalingam and Khera 2024, Ayaz-Alkaya and Kulakçı-Altıntaş 2025]. Embora ainda não sejam reconhecidos como condições clínicas nos manuais de diagnóstico oficiais [World Health Organization 2019, noa 2021], causam impactos nas funções executivas, como atenção e memória [Roswendi and Gong 2025], e na saúde mental, podendo agravar quadros de ansiedade, depressão, medo de ficar de fora (FoMO) e nomofobia [Ayaz-Alkaya and Kulakçı-Altıntaş 2025, Fang et al. 2026, Côrte-Real et al. 2022, Dam et al. 2023, Côrte-Real et al. 2022].

Entender tal fenômeno, portanto, é crucial para o desenvolvimento de intervenções. Contudo, a literatura em Tecnologia da Informação (TI) aplicada ao UPMS (Seção 2) concentra-se majoritariamente em classificadores supervisionados sobre dados autorreportados e comportamentais, muitas vezes sem sistematizar a relação entre os constructos teóricos que caracterizam o fenômeno e os sinais computacionalmente extraíveis a partir de dados públicos e não estruturados. Diante disso, o presente trabalho se norteia a partir do questionamento: **Como métodos computacionais e conceitos interdisciplinares podem ser sistematizados em um *framework* para a identificação e análise de padrões do UPMS?**

Para responder a essa pergunta, o objetivo geral deste trabalho é propor um *framework* para a sistematização do processo de identificação de padrões de UPMS integrando: a) uma camada teórica ancorada em construtos interdisciplinares precisos; b) uma camada de mapeamento que conecte tais constructos a sinais computacionalmente extraíveis de dados textuais ruidosos; e c) uma camada de implementação técnica, da qual o *pipeline* de extração e curadoria de dados do YouTube (Seção 3) constitui a primeira instanciação e prova de conceito. Para tanto, foram definidos os seguintes objetivos específicos:

- Sistematizar, por meio de revisão de literatura, os constructos teóricos que caracterizam o UPMS e as lacunas dos artefatos de TI existentes para sua identificação;
- Definir um modelo de mapeamento entre os constructos teóricos e sinais em textos de mídias sociais;
- Implementar e validar um *pipeline* de extração e filtragem semântica híbrida para dados ruidosos;
- Validar empiricamente o mapeamento por meio de anotação humana para aferir precisão;

- Avaliar a generalização do *framework* para outras plataformas sociais.

2. Trabalhos Relacionados

A partir de um mapeamento sistemático da literatura, identificaram-se as principais abordagens computacionais aplicadas à intervenção e detecção do UPMS. A literatura atual concentra-se no uso de classificadores supervisionados tradicionais aplicados a dados comportamentais e autorrelatados.

Nessa vertente, [Munia et al. 2025] alcançaram 97% de acurácia com GaussianNB para detecção precoce de transtornos, enquanto [Mim et al. 2024] e [Akter et al. 2022] usaram *Random Forest* e Regressão Logística com resultados semelhantes, alcançando altas acurácias (de 82% a 97%). Embora eficazes, essas abordagens compartilham a limitação da dependência de formulários manuais, os quais podem inserir viés de percepção.

Já [Yeasmin and Islam 2026], buscando explicabilidade, propuseram um *framework* preditivo fundamentado em uma escala clínica validada (*Bergen Social Media Addiction Scale* - BSMAS[1]). Utilizando o algoritmo *Gradient Boosting* aliado a técnicas de IA Explicável (SHAP e LIME) e inferência causal, os autores buscam mitigar o problema da ”caixa-preta”dos modelos tradicionais.

Apesar dos avanços, nota-se uma lacuna: a escassez de artefatos capazes de extrair sinais de UPMS de forma automatizada em grandes volumes de textos públicos (*User-Generated Content*). A complexidade semântica das mídias sociais desafia os modelos tradicionais, permitindo a exploração de abordagens híbridas de Processamento de Linguagem Natural (PLN) ancoradas na literatura.

3. Proposta

Para mitigar as limitações identificadas na literatura, o presente trabalho propõe um *framework* analítico estruturado em duas dimensões interdependentes: uma base conceitual para o mapeamento taxonômico e uma arquitetura tecnológica de extração e análise de dados ruidosos.

A primeira camada do *framework* consiste na correlação sistemática entre os constructos clínicos da dependência e os sinais textuais extraíveis. Para tanto, adotam-se os seis componentes de adicção de Griffiths (Saliência, Modificação do Humor, Tolerância, Abstinência, Conflito e Recaída) [Fang et al. 2026, Dam et al. 2023, Côrte-Real et al. 2022] como âncoras conceituais. A formalização estruturada desta matriz de mapeamento encontra-se em desenvolvimento e visa cruzar os dados empíricos extraídos pela camada tecnológica junto à literatura, como esforço para garantir um rigoroso embasamento da inferência computacional prevista.

Já na camada tecnológica, a viabilidade técnica do *framework* foi instanciada por meio de um *pipeline* desenvolvido em linguagem Python, projetado para a extração, préprocessamento e análise de dados não estruturados do YouTube. A metodologia do experimento divide-se em quatro fases: (a) extração, (b) higienização, (c) filtragem semântica, e (d) análise dos dados, detalhadas a seguir.

[1]https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft74607-000

A extração dos dados é realizada via *YouTube Data API v3*[2], cujos metadados são filtrados para descartar vídeos de curta duração e em idiomas diferentes do inglês. Após isso, é aplicada a higienização textual por meio de expressões regulares para remoção de ruídos absolutos (textos vazios ou compostos exclusivamente por *emojis*).

Já a filtragem semântica constitui o núcleo computacional do *pipeline*, uma abordagem híbrida projetada para lidar com a ambiguidade da linguagem natural. A triagem inicial emprega o modelo *Sentence Transformers* (ST) para calcular a similaridade de cosseno entre os comentários extraídos e frases-âncora definidas manualmente, automaticamente descartando ou aceitando comentários com baixa e alta similaridade, respectivamente. Os comentários na "área cinzenta", com pontuações fora dos intervalos de confiança, são encaminhados ao *Large Language Model* (LLM) *gemma-4-26b-a4b-it*[3], que realiza uma inferência contextual mais profunda para definir a aceitação ou rejeição final do comentário. Essa arquitetura atua como um filtro de alta performance, reduzindo o custo computacional e o volume de requisições de API, contornando limitações de cota.

Por fim, os comentários aceitos compõem o *corpus* utilizado para a análise de tópicos latentes por meio do modelo BERTopic[4] e análise de sentimentos por meio da biblioteca *Natural Language Toolkit*[5] com o módulo VADER[6]. Os resultados dessas análises foram cruzados analiticamente a fim de obter uma visão mais robusta e abrangente do fenômeno observado.

4. Resultados Parciais e Discussão

A arquitetura híbrida desenvolvida demonstrou eficiência e viabilidade técnica na mitigação de ambiguidades inerentes a textos de mídias sociais. A triagem inicial com ST atuou como filtro eficaz para os extremos de confiança, reduzindo drasticamente o custo computacional e o volume de requisições repassadas ao LLM. Essa estratégia otimizou o tempo de inferência global e contornou limitações de cota de API, reservando o processamento complexo apenas para os casos da área cinza.

A coleta via *YouTube Data API v3*, orientada por duas expressões de busca, retornou 148 vídeos e 83.025 comentários. Deste total, uma amostra de 38.911 comentários foi submetida ao *pipeline* híbrido, resultando na retenção de 3.127 textos válidos. Este corpus final subsidiou a extração automatizada de tópicos latentes, a análise de sentimentos e o cruzamento analítico entre ambas as frentes.

A modelagem de tópicos identificou 16 *clusters* válidos (excluindo-se o tópico de ruído -1), que foram categorizados em três macroeixos temáticos: 1) Consciência da inatenção e *doomscrolling*; 2) Papel dos algoritmos e indústria do engajamento; e 3) Táticas de contenção e saúde mental. O cruzamento destes eixos com a análise de sentimentos (Figura 1) evidenciou o teor emocional de cada *cluster*, revelando nuances críticas sobre a percepção da comunidade.

O primeiro eixo (Tópicos 0, 3, 7, 9 e 13) revela a alteração no padrão de consumo dos usuários, pautada pela intolerância ao tempo natural de absorção da informação. O

[2]https://developers.google.com/youtube/v3
[3]https://ai.google.dev/gemma/docs/core?hl=pt-br
[4]https://maartengr.github.io/BERTopic/index.html
[5]https://www.nltk.org/
[6]https://www.nltk.org/api/nltk.sentiment.vader.html

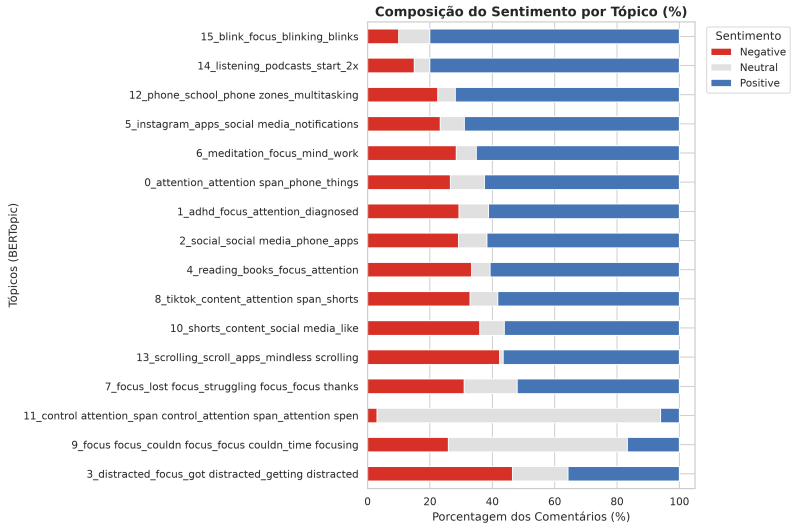

Figura 1. Análise entre os tópicos latentes e sentimentos predominantes

Tópico 0 (n=1.333), definido por termos como *attention span* e *speed*, ilustra a necessidade de acelerar vídeos (ex: “eu precisei colocar este vídeo em velocidade 2x para assistir a ele por completo e não rolar para outro vídeo”, traduzido do inglês) para mitigar a desatenção. De forma complementar, os tópicos 3, 7 e 9 (n≈400) reúnem confissões de falha cognitiva e perda de controle (*got distracted*, *lost focus*). Esse atrito culmina no Tópico 13 (n=84), que descreve o *doomscrolling* (*mindless scrolling*) como um estado de transe, refletindo a preocupação dos usuários com a automatização involuntária de seus comportamentos.

O segundo eixo (Tópicos 2, 5, 8 e 10; n=575) desloca o foco do déficit individual para a arquitetura das plataformas. Os usuários apontam o formato de vídeos curtos (*tiktok*, *shorts*) e algoritmos de recomendação (*recommendations*) como catalisadores da redução da capacidade de atenção. Os relatos expressam profunda frustração com gatilhos de engajamento (*notifications*, *screen time*) e frequentemente culminam em ações radicais de repúdio, como bloquear ou deletar os aplicativos (ex: “A caminho de deletar as redes sociais”, traduzido do inglês). A carga predominantemente negativa deste eixo sugere que as interfaces projetadas para a economia da atenção podem gerar o esgotamento emocional e levar a comunidade a perceber a inatenção como uma consequência sistêmica do *design*, e não apenas como falta de disciplina.

Por fim, o terceiro eixo (Tópicos 1, 4, 6, 12, 14 e 15; n=691) ilustra as consequências na saúde mental e as estratégias de mitigação. Frente à persuasão das interfaces, a comunidade adota mecanismos de readaptação cognitiva (*coping mechanisms*) por meio da migração para mídias de atenção sustentada (*reading*, *podcasts*) e de práticas de regulação fisiológica, como meditação (*meditation*, *blinking*). Tais *clusters* concentram a maior carga de sentimentos positivos, refletindo o alívio na retomada da agência sobre o próprio tempo. Em contrapartida, este eixo expõe a patologização da desatenção (Tópico 1; n=307). Dominado por termos como *adhd* e *diagnosed*, o *cluster* revela a dificuldade dos usuários em distinguir entre a exaustão cognitiva induzida pelo uso excessivo de telas (*brain fog*) e transtornos neurológicos clínicos, como o Transtorno de Déficit de Atenção e Hiperatividade (TDAH). Essa confusão sugere uma dificuldade da comunidade em dis-

tinguir exaustão cognitiva induzida pelo uso de telas de quadros clínicos, o que aponta para uma área de investigação futura.

5. Conclusão

Os resultados preliminares atestam a robustez do *pipeline* desenvolvido, evidenciando que a arquitetura híbrida de PLN é uma solução escalável para o tratamento de bases de dados sociais não estruturadas. A integração entre *Sentence Transformers* e o modelo LLM mostrou-se eficaz no tratamento da "área cinzenta", equilibrando custo computacional e precisão semântica para superar o ruído inerente aos textos de mídias sociais. Adicionalmente, a modelagem de tópicos mostra a viabilidade de se extrair agrupamentos semânticos aderentes aos constructos clínicos da psicologia, os critérios de Griffiths.

Do ponto de vista da área de Banco de Dados, uma das principais contribuições desta pesquisa reside na estruturação e enriquecimento semântico dos dados, com um *pipeline* desenhado para permitir a proveniência dos dados, viabilizando o rastreamento de todo o ciclo de vida da informação.

Como trabalhos futuros e próximos passos metodológicos, destacam-se: (1) a formalização estruturada da matriz de mapeamento taxonômico entre os dados do experimento e os critérios de Griffiths; (2) a validação deste mapeamento (3) a generalização da ingestão para outras APIs, comprovando sua interoperabilidade. Tais avanços visam amadurecer a prova de conceito atual em um artefato robusto de apoio à saúde digital. Além disso, é prevista a realização do exame de Qualificação de Mestrado entre julho e agosto de 2026.

6. Considerações

Declaração de Uso de Inteligência Artificial: Neste trabalho foram utilizados recursos de IA para o auxílio na revisão linguística (gramática, sintaxe e síntese). Os autores garantem a revisão e responsabilidade sobre todo o conteúdo.

Agradecimentos: O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - código de financiamento 001.

Referências

(2021). *Manual diagnóstico e estatístico e transtornos mentais: DSM-5*. Artmed.

Akter, M., Ritu, K. F., Habib, M. T., Rahman, M. S., and Ahmed, F. (2022). A Machine Learning Approach To Predict Social Media Addiction During COVID-19 Pandemic. In *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pages 401–405, Salem, India. IEEE.

Amirthalingam, J. and Khera, A. (2024). Understanding Social Media Addiction: A Deep Dive. *Cureus*.

Ayaz-Alkaya, S. and Kulakçı-Altıntaş, H. (2025). Social media addiction, nomophobia, and social anxiety among adolescents: A mediation analysis. *Journal of Pediatric Nursing*, 85:16–21.

Côrte-Real, B., Cordeiro, C., Câmara Pestana, P., Duarte E Silva, I., and Novais, F. (2022). Addictive Potential of Social Media: A Cross Sectional Study in Portugal. *Acta Médica Portuguesa*, 36(3):162–166. Portugal.

Dam, V. A. T., Dao, N. G., Nguyen, D. C., Vu, T. M. T., Boyer, L., Auquier, P., Fond, G., Ho, R. C. M., Ho, C. S. H., and Zhang, M. W. B. (2023). Quality of life and mental health of adolescents: Relationships with social media addiction, Fear of Missing out, and stress associated with neglect and negative reactions by online peers. *PLOS ONE*, 18(6):e0286766.

De La Torre, P. G., Pérez-Verdugo, M., and Barandiaran, X. E. (2025). Attention is all they need: cognitive science and the (techno)political economy of attention in humans and machines. *AI & SOCIETY*.

Desai, S. and Thombre, S. (2025). User Interactions and Attention as Internet Currency: Impact on Digital Assets. In *2025 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pages 1–6, Bengaluru, India. IEEE.

Fang, X., Yang, J., Liu, B., Wang, Y., Liao, J., and Lei, L. (2026). Exploring fear of missing out and problematic social media use among Generation Alpha: a random intercept cross-lagged panel model and cross-lagged panel network analysis. *Addictive Behaviors*, 172:108489.

IBGE, editor (2025). *Acesso à internet e à televisão e posse de telefone móvel celular para uso pessoal 2024*. IBGE, Rio de Janeiro, RJ.

Mim, M. N., Firoz, M., Islam, M. M., Hasan, M., and Habib, M. T. (2024). A study on social media addiction analysis on the people of Bangladesh using machine learning algorithms. *Bulletin of Electrical Engineering and Informatics*, 13(5):3493–3502.

Munia, J. A., Tahmiduzzaman, K., Saad, I. H., Islam, M. M., Akash, A. H., and Zumma, M. T. (2025). A Study of Predicting Social Networking Sites(SNS) Media Disorder Survey Data Using a Machine Learning Approach. In *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, pages 300–306, Erode, India. IEEE.

Roswendi, R. and Gong, X. (2025). A Cross-Cultural Study on the Impact of Social Media Usage on High School Students' Learning. *IEEE Integrated STEM Education Conference*.

World Health Organization (2019). International statistical classification of diseases and related health problems (11th ed.). `https://icd.who.int/browse/2025-01/mms/pt`. Acessado em: 14 maio 2026.

Yeasmin, T. and Islam, M. M. (2026). A Data-Driven Approach to Combating Social Media Addiction: Insights from Explainable AI. In *2026 5th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, pages 1–6, Rajshahi, Bangladesh. IEEE.