



Inteligência artificial como tecnologia para guerras híbridas

RESUMO

Conflitos híbridos combinam métodos convencionais e não convencionais, incluindo desinformação para influenciar a opinião pública e desestabilizar sociedades, conforme [6]. Este trabalho desenvolve uma tecnologia inédita no âmbito da defesa nacional através de um MVP (mínimo produto viável) do tipo protótipo, visando a análise do ambiente informacional em conflitos híbridos modernos, especialmente no ciberespaço. O grande volume e a diversidade de dados tornam o ambiente desafiador. O presente trabalho destaca o emprego de requisitos funcionais, como Big Data, machine learning e bases de dados distribuídas, estabelecidos por [4]. Uma aplicação web serve como interface, onde é informada uma URL para web scraping. Os dados coletados são armazenados em uma base de dados NoSQL e processados por uma API que utiliza uma rede neural convolucional para classificar a notícia e analisar a polaridade (sentimento) do texto. A pesquisa justifica-se pela necessidade de tecnologias baseadas em IA (inteligência artificial) para enfrentar a desinformação e criar estratégias de defesa eficazes, conforme [4]. O objetivo é fortalecer a resiliência cognitiva e promover decisões acuradas em contextos complexos. A metodologia inclui o treinamento de uma rede neural convolucional com notícias da operação GLO [1]. O MVP provou ser útil na análise do ambiente informacional e na melhoria da consciência situacional, apesar de limitações na acurácia devido ao pequeno volume de dados históricos de outras operações.

PALAVRAS-CHAVE: IA. Big Data. Desinformação. Guerras Híbridas. MVP.

1 INTRODUÇÃO

Daniel Kahneman, Nobel de Economia em 2002, revela que os humanos tendem a superestimar seu autocontrole e racionalidade, o que afeta a objetividade na tomada de decisões. Em [2], Kahneman descreve dois sistemas mentais: o Sistema 1, rápido e intuitivo, e o Sistema 2, mais lento e analítico. A mente frequentemente evita o esforço do Sistema 2, o que pode ser manipulado por estereótipos e desinformação, prejudicando a análise crítica. Esse fenômeno é crítico em conflitos híbridos, onde a desinformação explora essas vulnerabilidades cognitivas para manipular percepções e comportamentos, há ainda o fato observado em [3], onde as notícias falsas alcançam mais pessoas e se difundem mais rapidamente, que notícias verdadeiras. Percebendo a pluralidade e natureza do terreno que é o ciberespaço, a velocidade de propagação das informações e natureza dos dados, o MVP é um produto que têm por objetivo ser uma ferramenta de análise do ambiente informacional conforme requisitos de [4].

2 METODOLOGIA

Adotou-se a metodologia Design Thinking [5], para o desenvolvimento de um MVP do tipo protótipo. Dessa forma o projeto foi estruturado em duas dimensões, no contexto do



problema e no contexto da solução. O Contexto do problema contém um subdomínio de entendimento que têm uma natureza de ideias divergentes, onde foi compreendido as regras de negocio, requisitos funcionais observados no campo da operação GLO [1] e na publicação [4]. A fase seguinte é a etapa final do contexto do problema possui a etapa de convergência, o subdomínio de definição, onde foi definido um objetivo SMART (especifico, mensurável, atingível, relevante e temporal): ter um produto que pudesse ser utilizado em até 4 meses na versão beta com inteligência artificial para análise do ambiente informacional (figura 01).



FIGURA 01:

CANVAS DO MVP. FONTE: PRÓPRIO AUTOR.

A FASE DE DEFINIÇÃO INTERSECTA COM O CONTEXTO DA SOLUÇÃO, CUJA ETAPA INICIAL, CHAMADA DE DESENVOLVIMENTO, É O SUBDOMÍNIO DE CONVERGÊNCIA. NESTA FASE, FOI SELECIONADO UM CONJUNTO DE POSSÍVEIS FERRAMENTAS E FRAMEWORKS QUE ATENDIAM AOS REQUISITOS FUNCIONAIS DO OBJETIVO SMART. PARA ISSO, FOI CRIADO UM AMBIENTE DE EXPERIMENTAÇÃO COM JUPYTER NOTEBOOK, DOCKER, VIRTUALENV, PYTHON, STREAMLIT, MONGODB, MONGO EXPRESS, TENSORFLOW E DJANGO. APÓS O DESENVOLVIMENTO, INICIOU-SE A FASE DE CONSTRUÇÃO DO MVP DO TIPO PROTÓTIPO, SEGUINDO UMA ARQUITETURA DE SOFTWARE ORIENTADA À FILOSOFIA SOA (ARQUITETURA ORIENTADA A SERVIÇOS). ESTA ARQUITETURA FOI DIVIDIDA EM QUATRO APLICAÇÕES:

- Três aplicações foram containerizadas usando Docker, com as imagens MongoDB, Mongo Express e a aplicação desenvolvida, nomeada de app_01.
- A quarta aplicação, app_02, foi desenvolvida em um ambiente virtual local utilizando VirtualEnv, isolando-a de outros softwares.

A aplicação app_01 fornece uma interface web construída com o framework Streamlit de Python. A aplicação app_02 recebe requisições via API, construída com Django, um framework web de Python, e executa rotinas responsáveis pela classificação. Para a classificação, utilizou-se TensorFlow para construir a rede neural convolucional, e a persistência dos dados é feita na base de dados NoSQL MongoDB, que pode funcionar de forma distribuída e ser gerenciada via web com o Mongo Express. A



explicação e visualização do MVP estão disponíveis no YouTube e GitHub em [7] e [8], respectivamente.

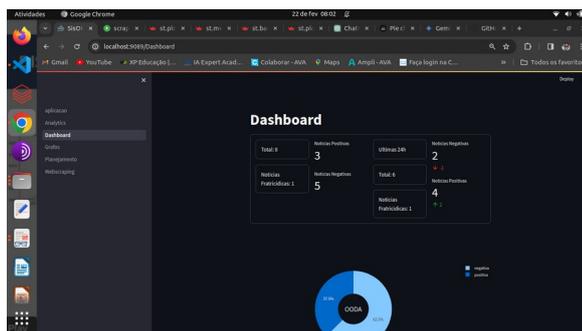


FIGURA 02 – PRINTSCREEN DA TELA DE DASHBOARD DO MVP. FONTE:PRÓPRIO AUTOR.

Etapa 1 Coleta de dados via Webscraping:

Os dados foram obtidos de fontes públicas na internet e o web scraping foi realizado utilizando as bibliotecas **requests** e **BeautifulSoup** do Python. A biblioteca **requests** foi usada para fazer as requisições HTTP às páginas da web, enquanto **BeautifulSoup** foi empregada para a análise e extração dos dados HTML das páginas. É importante ressaltar que, como HTML é um tipo de dado semiestruturado, podem ocorrer mudanças nas metatags das páginas, o que pode requerer a atualização do XPATH utilizado na extração dos dados.

Etapa 2 Data Understanding e Data Preparation :

Na segunda etapa, que envolve a compreensão e preparação dos dados, a análise exploratória, o pré-processamento e a limpeza dos dados foram realizados no **JupyterLab**. Neste ambiente de experimentação, foram executados procedimentos fundamentais, incluindo a limpeza dos textos, remoção de stop words, tokenização e criação de embeddings. Esses processos foram essenciais para preparar os dados antes da definição e construção das classes finais. Além disso, a modelagem da rede neural convolucional para classificar a polaridade dos textos como positiva ou negativa também foi realizada no JupyterLab. Todo o processo seguiu o framework CRISP-DM (Cross Industry Standard Process for Data Mining), garantindo uma abordagem estruturada e sistemática para a análise e modelagem dos dados.

Etapa 3 Construção das Classes e Arquitetura Orientada a Serviços



Nesta etapa, focou-se na construção das classes seguindo as boas práticas de design de software, com ênfase na redução de acoplamento e na criação de uma arquitetura orientada a serviços (SOA). Os modelos testados com dados pré-processados do web scraping passaram por uma fase de modelagem orientada a objeto e funcional, para garantir uma estrutura robusta e modular. Parte da aplicação foi implementada no **app_02**, que é responsável pela API desenvolvida em Django. Esta API invoca o classificador baseado em rede neural convolucional para analisar e classificar o texto. Após a classificação, os resultados são armazenados na base de dados MongoDB. O status da operação é retornado para o usuário via API Django. O **app_01**, desenvolvido com Streamlit e operando em um ambiente virtual, é responsável por renderizar os dados e resultados retornados pela API do **app_02**. A integração entre esses componentes assegura uma comunicação eficiente e a apresentação dos resultados de forma interativa e acessível.

Etapa 4 Desenvolvimento da Rede Neural Convolucional (CNN):

Nesta etapa, desenvolveu-se a Rede Neural Convolucional (CNN) utilizando os dados preprocessados armazenados na base de dados MongoDB. O embedding do corpus foi realizado com a técnica CBOW (Continuous Bag of Words), utilizando a biblioteca TensorFlow. A escolha da arquitetura CNN com foi motivada pela necessidade de um modelo que requer menos dados para treinamento em comparação com Modelos de Linguagem de Grande (LLMs) e que apresenta uma eficiência significativa na tarefa de classificação, além de ser menos custoso computacionalmente do que os Transformers. Essa escolha se deu devido à limitação de poder computacional disponível e à quantidade inicial de dados, que era relativamente pequena. Assim, a CNN proporcionou um equilíbrio entre desempenho e eficiência, adequado para o contexto e os recursos disponíveis.

3 RESULTADOS E DISCUSSÃO

O MVP (Produto Mínimo Viável) do tipo protótipo demonstrou ser uma prova de conceito valiosa, desenvolvida com base nos principais requisitos estabelecidos pelo [4]. O projeto visou mitigar a paralisia da informação causada pelo grande volume de dados, permitindo a manutenção eficiente do ciclo OODA (Observar, Orientar, Decidir e Agir). Isso contribuiu para uma consciência situacional mais rápida e precisa do ambiente informacional, promovendo uma cultura militar orientada por dados. O produto foi apresentado ao CoNavOpesp, que decidiu avançar no desenvolvimento, buscando criar uma aplicação mais robusta e sofisticada. Em julho de 2024, o projeto está em fase de redesign da arquitetura de Machine Learning. A nova abordagem inclui a integração de modelos de Linguagem de Grande Escala (LLMs) através de uma API utilizando LangChain, com suporte para APIs do ChatGPT e Gemini, empregando a técnica RAG (Retrieval-Augmented Generation) para a classificação precisa de notícias. Esse redesign visa eliminar a necessidade de um parque tecnológico próprio e permitirá a utilização dos dados classificados conforme uma engenharia de prompt pré-determinada, ajustada para diferentes objetivos de operação da informação. Futuramente, os dados classificados serão utilizados para treinar uma rede neural própria, que, a princípio, será



baseada em Transformers. Dependendo dos resultados obtidos, o treinamento poderá envolver técnicas de fine-tuning com LPM (modelos de linguagem pré-treinadas) como BERTimbau ou modelos LLaMA 2 ou 3.

4 CONSIDERAÇÕES FINAIS

O projeto proporciona ganhos significativos para a gestão do conhecimento ao utilizar tecnologias de big data, armazenamento de dados estruturados e não estruturados, e inteligência artificial. Essas tecnologias permitem uma rápida consciência situacional do campo informacional, favorecendo uma tomada de decisão mais assertiva e baseada em dados históricos. A solução garante também a possibilidade de auditoria, pois armazena e documenta todos os dados e raciais utilizados. Dessa forma, assegura-se que as decisões são informadas e rastreáveis, otimizando a gestão da informação e a resposta a situações complexas.

REFERÊNCIAS

1. BRASIL. Decreto nº 11.765, de 15 de novembro de 2023. Autoriza o emprego das Forças Armadas para a Garantia da Lei e da Ordem em portos e aeroportos. Diário Oficial da União: seção 1, Brasília, DF, 15 nov. 2023.
2. Kahneman, D. (2012). **Rápido e Devagar: Duas Formas de Pensar**. Rio de Janeiro: Objetiva.
3. Vosoughi S, Roy D, Aral S (2018a) A disseminação de notícias verdadeiras e falsas online. *Sci* 359:1146–1151. <https://www.science.org/doi/abs/10.1126/science.aap9559> , <https://doi.org/10.1126/science.aap9559>
4. ESTADO-MAIOR DA ARMADA (Brasil). **EMA-335 Doutrina de Operações de Informação** . Brasília: EMA, 2018. 62 p.
5. Brown, T. (2020). **Design Thinking: Uma metodologia poderosa para decretar o fim das velhas ideias (Edição Comemorativa 10 Anos)**. Rio de Janeiro: Alta Books.
6. Escola de Guerra Naval. **Tese apresentada como requisito parcial para a conclusão do Curso de Política e Estratégia Marítimas da Escola de Guerra Naval. Produzido por CMG (FN) Alexandre Henrique Batista Barbosa**. Disponível em: <https://www.marinha.mil.br/egn/sites/www.marinha.mil.br/egn/files/C-PEM010%20-%20CMG%20%28FN%29%20ALEXANDRE%20HENRIQUE%20BATISTA%20BARBOSA%20-%20A%20DESINFORMA%C3%87%C3%83O%20COMO%20FERRAMENTA%20DA%20GUERRA%20H%C3%8DBRIDA.pdf>. Acesso em: 12 jul. 2024, às 12h50.
7. PESSOA, Alessandro. **Pitch MBA Ciência de Dados - Sistema Operação Da Informação Aplicativo**. Disponível em: <https://youtu.be/yyDdlp2uo4M?si=RSw1cYhEi-fHAXGT>. Acesso em: 3 ago. 2024.



SIMPÓSIO DE GESTÃO DO CONHECIMENTO DA MB

DESAFIOS E OPORTUNIDADES
NA ERA DIGITAL
27 E 28
NOVEMBRO DE 2024

LOCAL: AUDITÓRIO AMAZÔNIA AZUL - ESCOLA NAVAL

-
8. Pessoa, Alessandro. (2024). *Projeto Aplicado em Sistemas Operacionais e Informativos* [repositório GitHub]. Disponível em <https://github.com/alessandropessoa/projetoAplicadoSisOpInfo>