



## **DISRUPTIVE ARCHITECTURES: IOT, IOB & GENERATIVE IA**

Enzo Ross Gallone – 551754 - 2TDSPM

Gabriel de Andrade Baltazar - 550870 – 2TDSPM

Leonardo Cordeiro Scotti – 550769 - 2TDSPM

Pedro Gomes Fernandes - 551480 - 2TDSPM

Vinícius de Araújo Camargo - 99494 - 2TDSPM

São Paulo, SP

2024

# **Documentação do Projeto: Previsão de Tendências no Mercado de Tecnologia com IA Generativa**

Projeto de previsão de tendências no mercado de tecnologia, focado em tópicos de Inteligência Artificial (IA) e Machine Learning (ML). Utilizando uma rede neural LSTM (Long Short-Term Memory), a solução visa identificar padrões de popularidade ao longo do tempo, fornecendo insights valiosos para o setor tecnológico.

---

## **Resumo do Projeto**

O projeto utiliza uma rede neural LSTM para prever tendências de popularidade de IA e ML. A LSTM foi escolhida pela sua capacidade de processar dados sequenciais e capturar dependências de longo prazo, sendo ideal para dados de séries temporais. Com um conjunto de dados de popularidade, o modelo foi treinado e avaliado para prever mudanças de interesse no setor, permitindo antecipar picos e quedas na popularidade dos tópicos.

## **Arquitetura LSTM**

A arquitetura de rede LSTM foi projetada para:

- **Captura de Padrões Temporais:** As unidades de memória da LSTM mantêm informações ao longo do tempo, o que ajuda a identificar tendências duradouras.
- **Portas de Controle Inteligentes:** A LSTM possui portas de entrada, esquecimento e saída, que gerenciam a retenção e descarte de informações, evitando o desvanecimento do gradiente e melhorando a eficácia com dados de longa duração.

## **Estrutura e Implementação do Modelo**

O modelo foi estruturado em três principais componentes:

1. **Camadas LSTM:** Duas camadas LSTM sequenciais foram implementadas. A primeira camada processa a sequência completa, e a segunda extrai características temporais detalhadas.

2. **Camada Densa de Saída:** Uma camada densa com uma unidade de saída gera a previsão final da popularidade do tópico.
- 

## **Treinamento e Avaliação do Modelo**

Para otimizar o treinamento, algumas técnicas foram aplicadas:

- **Normalização dos Dados:** O `MinMaxScaler` foi usado para ajustar os dados ao intervalo  $[0, 1]$ , melhorando a eficiência do modelo.
- **Configuração de Janelas Temporais:** Os dados foram organizados em blocos temporais para que o modelo pudesse capturar dependências com observações passadas.
- **Parâmetros de Treinamento:** O otimizador `adam` e a função de perda `mean_squared_error` foram usados para ajustar o modelo, com o objetivo de minimizar o erro.

## **Métricas de Avaliação**

Para avaliar o desempenho em dados de teste, foram usadas métricas como:

- **MSE (Erro Quadrático Médio):** Mede a média dos erros quadrados entre previsão e valor real.
  - **R<sup>2</sup> Score:** Avalia o quanto a variação nos dados de teste é explicada pelo modelo.
  - **MAE (Erro Médio Absoluto):** Indica a média dos erros absolutos.
  - **MAPE (Erro Médio Percentual Absoluto):** Calcula a média do erro relativo em porcentagem.
- 

## **Base de Dados e Pré-processamento**

O conjunto de dados contém informações sobre popularidade, pesquisas crescentes e tendências em IA e ML.

## **Etapas de Pré-processamento**

- 1. Tratamento de Valores Nulos:** Preenchimento com valores anteriores para garantir consistência.
- 2. Transformação de Variáveis Categóricas:** Aplicado one-hot encoding para variáveis categóricas, permitindo que o modelo intérprete esses dados.
- 3. Normalização:** Todos os dados foram normalizados para melhorar a convergência e o desempenho do modelo.

## **Divisão dos Dados**

Os dados foram divididos em conjuntos de treino e teste, para garantir uma avaliação justa do desempenho em dados não vistos.

---

## **Visualização e Insights**

A partir das previsões, visualizações foram criadas para destacar padrões de popularidade:

- **Mapa de Calor de Correlação:** Exibe as relações entre variáveis, ajudando a entender como diferentes fatores influenciam a popularidade.
  - **Gráfico de Previsões e Valores Reais:** Permite comparar previsões com valores reais, facilitando a análise da precisão do modelo.
  - **Identificação de Picos e Quedas:** Picos e quedas de popularidade foram destacadas, oferecendo insights sobre os períodos de maior e menor interesse.
- 

## **Análise Final e Reflexão**

O modelo LSTM mostrou-se eficaz em prever tendências e identificar variações de popularidade. Essa abordagem permitiu capturar padrões temporais importantes, fornecendo um método preciso para previsão de tendências no setor de IA e ML.

## **Reflexão sobre o Processo de Desenvolvimento**

- **Aspectos Positivos:** A escolha da LSTM e o pré-processamento dos dados foram bem-sucedidos, resultando em previsões satisfatórias.

- **Desafios e Melhorias:** O ajuste de hiperparâmetros foi um desafio, e uma base de dados mais extensa poderia melhorar a precisão do modelo.
  - **Aprendizados:** Foi possível aprender sobre a importância de uma preparação de dados robusta e de técnicas de avaliação para prever dados temporais.
- 

## **Planos para o Futuro do Projeto**

1. **Expansão da Base de Dados:** Adicionar mais fontes de dados para melhorar a robustez e a precisão.
2. **Experimentação com Modelos Híbridos:** Integrar modelos LSTM com outras arquiteturas, como redes neurais convolucionais (CNN), para capturar padrões mais complexos.
3. **Automatização de Ajustes de Hiperparâmetros:** Implementar uma busca automatizada de hiperparâmetros para otimizar o desempenho.
4. **Integração com Dashboards Interativos:** Criar visualizações interativas para facilitar a análise de tendências e previsões.