# **Melhores Práticas para a Criação de uma IA de Identificação de Golpes na Internet Integrada com WhatsApp**

## **1. Introdução**

Este documento detalha as melhores práticas para o desenvolvimento de uma Inteligência Artificial (IA) focada em **auxiliar usuários a identificar golpes na internet**, integrando-se ao WhatsApp para fornecer um canal de comunicação acessível e automatizado.

A Wex analisará mensagens suspeitas enviadas pelos usuários, detectando fraudes conhecidas e oferecendo recomendações preventivas realizando a análise das evidencias fornecidas.

## **2. Objetivo**

Criar um **chatbot inteligente** que:

* Analise mensagens suspeitas e identifique golpes virtuais.
* Forneça feedback instantâneo ao usuário.
* Aprimore continuamente sua base de conhecimento com aprendizado de máquina.

## **3. Escopo do Projeto**

### **Diferentes Contextos de Uso**

A IA será utilizada em múltiplos canais e dispositivos:

* **WhatsApp**: Suporte direto ao usuário.
* **Web**: Interface para acompanhamento de estatísticas e feedback.

### **Camadas da Interface**

A IA será estruturada em **três camadas principais** para garantir eficiência:

1. **Entrada (Input do Usuário)**: Usuário envia mensagens suspeita via WhatsApp ou Web
2. **Processamento (****Processamento de Linguagem Natural + Machine Learning)** : IA analisa a mensagem, verifica padrões e calcula um score de risco.
3. **Saída (Resposta e Feedback)**: A IA responde ao usuário com diagnóstico e recomendações.

## **4. Arquitetura e Tecnologias**

### **Backend**

* **Linguagem**: Node.js (com Express)
* **Banco de Dados**: DynamoDB (AWS)
* **Hospedagem**: AWS Lambda (serverless)
* **Integração WhatsApp**: API do WhatsApp Business, via Meta
* **Machine Learning**: TensorFlow.js

### **Processamento de Linguagem Natural (NLP)**

* **TensorFlow.js** : Treinamento e inferência do modelo diretamente no backend.
* **Geração de tokens & Análise Semântica** : Identificação de padrões em mensagens.
* **Pontuação de Risco** : Atribuição de nota baseada em características do golpe.

## **5. Treinamento da IA**

### **5.1 Base de Treinamento**

* **Mensagens fraudulentas reais**: Coletadas e classificadas.
* **Mensagens legítimas**: Para evitar falsos positivos.
* **Processo de limpeza e normalização**: Remoção de ruído nos dados.

### **5.2 Modelagem do Algoritmo**

**Classificação Binária**: Golpe ou Não Golpe.  
**Análise de Contexto** : Algoritmo de similaridade semântica.  
**Pontuação de Risco** : Nota de confiabilidade atribuída a cada mensagem.

## **6. Fluxo de Processo**

**O Usuário envia mensagens suspeita via WhatsApp ou Web.**  
**A IA analisa o texto e extrai palavras-chave.**  
**O Modelo compara com base de dados dos golpes conhecidos.**

**A IA responde ao usuário com diagnóstico e sugestões**  
**O Usuário fornece feedback para melhorar o modelo**

## **7. Integração com WhatsApp**

### **7.1 Uso da API do WhatsApp Business**

* Configuração via Facebook Developer.
* Uso de **Webhooks** para receber mensagens.
* Respostas automáticas geradas pela IA.

### **7.2 Armazenamento de Dados**

* **Registro de interações:** Para aprimoramento contínuo da IA.
* **Logs anonimizados:** Proteção de privacidade.
* **Armazenamento** :DynamoDB com alta escalabilidade

## **8. Requisitos do Sistema**

### **Requisitos Funcionais (RF)**

1. **RF01 – Processamento de Linguagem Natural**
   * A IA deve interpretar mensagens usando a Processamento de Linguagem Natural para identificar padrões de golpes.
2. **RF02 – Classificação de Golpes**
   * O modelo deve classificar mensagens como **Golpe ou Não Golpe** com um score de risco.
3. **RF03 – Feedback e Aprendizado Contínuo**
   * O usuário pode confirmar se a resposta foi útil, ajudando a IA a melhorar.

### **Requisitos Não Funcionais (RNF)**

1. **RNF01 – Tempo de Resposta**
   * A IA deve fornecer uma resposta em **menos de 2 segundos** para não comprometer a experiência do usuário.
2. **RNF02 – Segurança e Privacidade**
   * As mensagens devem ser anônimas antes de serem armazenadas para análise posterior.
3. **RNF03 – Escalabilidade**
   * O sistema deve ser capaz de processar **milhares de mensagens simultaneamente**, garantindo performance.

## **9. Melhorias Futuras**

* **Suporte a áudio e imagens** (detecção de fraudes visuais).
* **Expansão para novos canais** (Telegram, Messenger, e-mail).
* **Uso de IA Generativa** para responder dúvidas complexas sobre golpes.

## **Conclusão**

Esta IA proporcionará **mais segurança para os usuários** ao identificar golpes na internet, utilizando **machine learning e NLP para análise de mensagens suspeitas**. A **integração com WhatsApp** permite um contato direto com os usuários, fornecendo **respostas rápidas e confiáveis** para evitar fraudes.