**Relatório Técnico – Trabalho 2: Inteligência Artificial**

## 1. Definição da Abordagem

**Proposta escolhida (Opção 1):** Criação de modelo com dataset público para classificação de sentimento.

**Objetivo:** Desenvolver um classificador de sentimento para reviews de games, prevendo se a avaliação do usuário é positiva ou negativa. Este problema é relevante para análise de opiniões em plataformas de jogos e moderação de conteúdo.

**Justificativa:** Utilização de rede neural recorrente (LSTM) aplicada a texto tokenizado, demonstrando conceitos de pré‑processamento, arquitetura e validação de IA.

## 2. Preparação dos Dados

**Fonte do dataset:**

* Game Review Dataset disponível em Kaggle: <https://www.kaggle.com/datasets/arashnic/game-review-dataset?resource=download>

**Descrição:**

* **Amostras:** ~1.000 reviews de games: coluna user\_review, coluna user\_suggestion: 0 = negativo, 1 = positivo.

**Etapas de pré‑processamento:**

1. **Leitura dos dados:** train.csv lido via pandas.
2. **Limpeza de texto:** remoção de tags HTML e caracteres não alfabéticos; conversão para minúsculas.
3. **Tokenização:** restrição a 10.000 palavras mais frequentes com Tokenizer do Keras.
4. **Sequenciamento:** conversão de cada review em sequência de índices numéricos com texts\_to\_sequences.
5. **Padding:** padronização do comprimento para 200 tokens com pad\_sequences.

## 3. Construção do Modelo

**Framework:** TensorFlow/Keras

* **Função de perda:** binary\_crossentropy
* **Otimizador:** Adam
* **Métrica:** acurácia

## 4. Teste, Validação e Análise Crítica

**Treinamento:**

* Batch size: 512
* Callback de EarlyStopping monitorando val\_loss (paciência=3)
* Dimensão do embedding: 16
* Unidades LSTM: 16
* Épocas: até 10

**Resultados no conjunto completo (treino + validação):**

* **Loss:** 0.3092
* **Acurácia:** 0.8776%
* **Recall:** 0.8739%
* **F1-score:** 0.8905%

## 5. Ajustes e Refinamento

Para refinar, testou-se:

* Dimensão do embedding: 16 vs. 64
* Unidades LSTM: 16 vs. 32

**Resultados:**

* **Loss:** 0.2647
* **Acurácia:** 0.9030%
* **Recall:** 0.9183%
* **F1-score:** 0.9152%

## 6. Conclusão

Este trabalho demonstrou a criação de um classificador de sentimento para avaliações de games, abrangendo todo o pipeline de IA: seleção e limpeza de dados, modelagem com LSTM, avaliação e refinamento.

## 7. Referências

1. Kaggle – Game Review Dataset: <https://www.kaggle.com/datasets/arashnic/game-review-dataset?resource=download>