**Estudo de Caso – Classificação de texto dos dados da John Deere utilizando LSTM e TRANSFORMER**

**Grupo:**

André Gomes Monteiro RM: 89168

Leonardo Aranha RM: 86919

Luara Maria Marino RM: 89375

Renato Kenji Yamashiro RM: 88847

**Resumo**

Numa empresa, a automação é fundamental para a eficiência de uma empresa, onde com esta ferramenta é possível a diminuição do erro humano, otimização do tempo e gerenciamento. Este estudo aborda o estudo de caso onde a John Deere deseja aumentar a produtividade encaminhando as mensagens para o setor correspondente.

**Introdução**

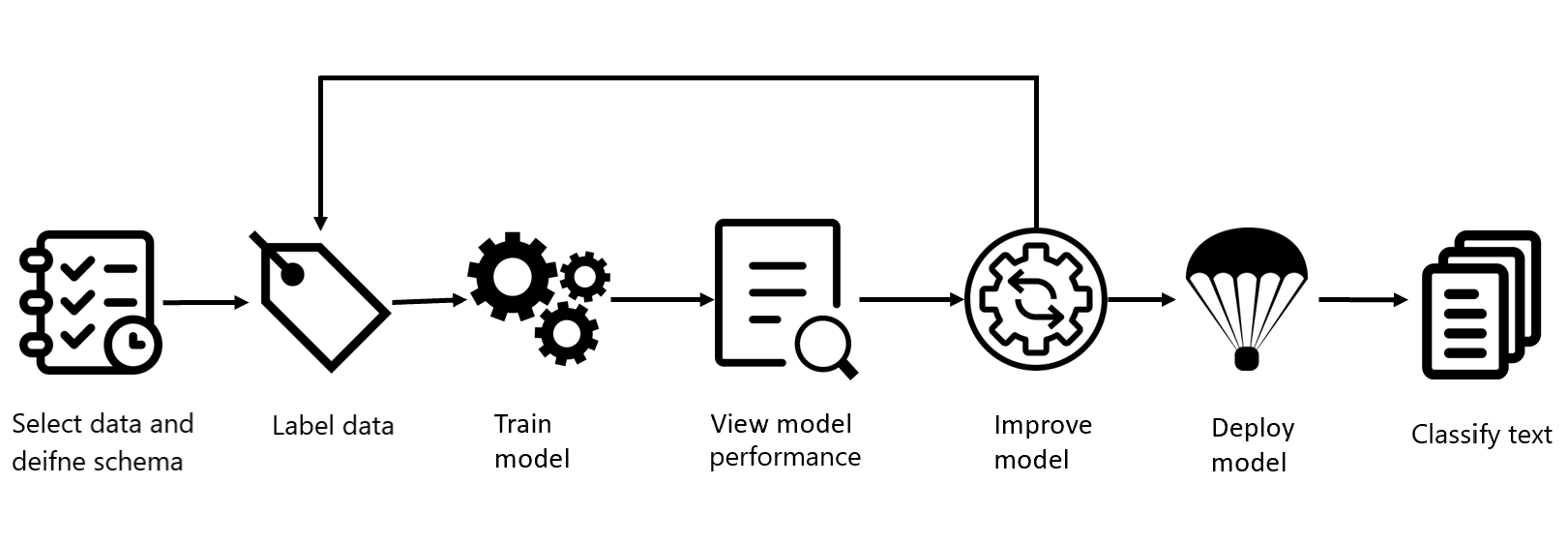
O processamento de linguagem natural usa machine learning para revelar a estrutura e o significado do texto. Com aplicativos de processamento de linguagem natural, as organizações podem analisar textos e extrair informações sobre pessoas, lugares e eventos. Este documento tem como finalidade analisar as possibilidades desta ferramenta para a otimização dos processos da John Deere e analisar os modelos LSTM e Transformer para classificação das mensagens.

**Metodologia**

**Arquitetura proposta**

Supomos que o direcionamento dos dados sejam feitos manualmente, utilizando o processamento de linguagem natural e machine learning, podemos identificar em que setor cada mensagem é encaminhada.

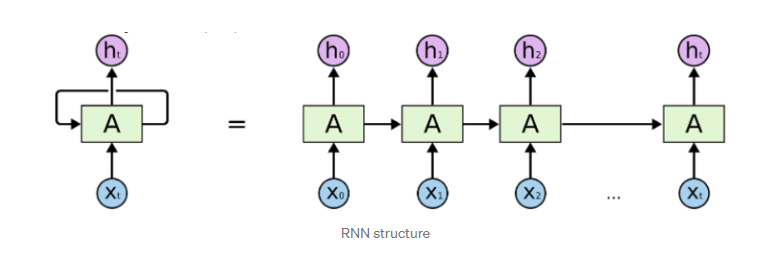
A imagem abaixo mostra a arquitetura proposta:



Analisando o dataset, encontramos so dados ‘u\_issue\_type’, 'category' e ‘subcategory’ como as possíveis targets no entanto, ao concatená-los, o número possível de targets aumentou para 42 classes, mantendo apenas ‘u\_issue\_type’ com 10 categorias.

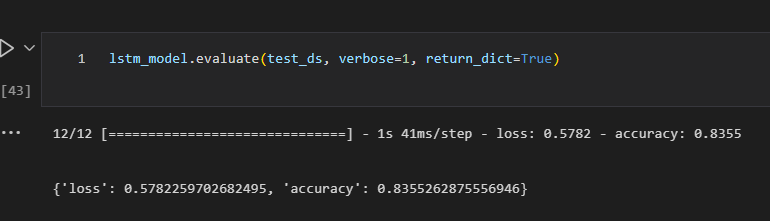
**LSTM**

LSTM é uma arquitetura de rede neural recorrente (RNN - acrônimo em inglês para recurrent neural network) que lembra os valores sobre um intervalo arbitrário. LSTM é usado para classificar textos e predição de séries temporais.



A arquitetura RNN por levar em conta a sequência dos pontos dos dados, leva em conta os estados ou as informações das entradas por meio de uma ‘memória’.

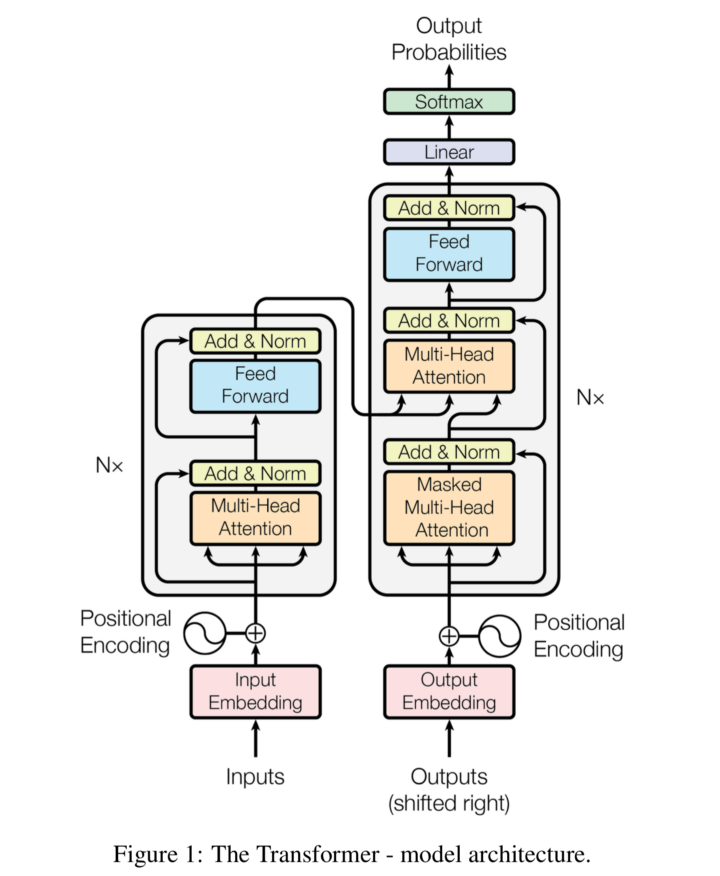
Os resultados obtidos para os dados de treino com a LSTM foi a acurácia de 83,5% com 50 épocas.



**Transformer**

Segundo a bibliografia ‘Attention Is All You Need’ descreve o transformer e a arquitetura ‘sequence-to-sequence’. Ela consiste num Encoder e um Decoder, onde o Encoder pega a sequência imputada e mapeia para uma outra dimensão (dimensão com n vetores) onde em seguida o Decoder solta uma saída em sequência. O resultado de saída pode ser outra língua, símbolos, cópia do output etc.

A característica principal do Transformer é o mecanismo de atenção onde analisa a sequência imputada e decide qual passo é importante para cada parte.



O modelo Transformer compilado com otimizador adam, perda categorical cross entropy e treinado com 20 épocas obteve acurácia nos dados de teste de 83,6%.

**Conclusão**

Analisando as métricas de acurácia, acreditamos que é possível utilizar Deep Learning para a classificação das mensagens da John Deere.

Como o dataset obtido não foi grande, é possível aumentar a acurácia do modelo.