



Geração automática de questões para a área de educação

João Victor, Maely Souza, Matheus Felipe





INTRODUÇÃO

Faremos uma **análise** do que recuperamos de **informação** no artigo **A Systematic Review of Automatic Question Generation for Educational Purposes** e do nosso andamento no **projeto**

Sobre nosso Primeiro Banco

Nesse projeto iniciamos tentando usar o **banco de dados** <https://allenai.org/data/sciq> . Segue uma breve descrição do que encontramos nele

- ❑ **Question:** Campo contendo a **questão**
- ❑ **Distractors:** Campo contendo as **opções incorretas** de respostas
- ❑ **Correct Answer:** Campo contendo a **opção correta** de resposta
- ❑ **Support:** Campo contendo um **texto** que **ajuda** no **entendimento** da resposta correta

O **Banco** está no formato **.json** e já possui uma divisão de **treino** e **teste**

As **perguntas** são na área de **ciências** e vem desse [paper](#)

Sobre nosso Segundo Banco

Após nossa primeira tentativa retornar resultados abaixo do esperado, trocamos para o **banco de dados** <https://rajpurkar.github.io/SQuAD-explorer/>. Segue uma breve descrição dos seus principais encontramos nele

- ❑ **Question:** Campo contendo a **questão**
- ❑ **Answer:** Campo contendo a **opção correta** de resposta
- ❑ **Context:** Campo contendo um **texto** que **ajuda** no **entendimento** da resposta correta

O **Banco** está no formato **.json** e já possui uma divisão de **teste** e **dev**

As **perguntas** são na área de **variedades** e são baseadas em textos da Wiki



Geração de Perguntas



? Problema

O problema em questão é a automática geração de perguntas a partir de um texto de apoio



Método



Para criar as questões seguimos os seguintes passos a passos:

- Analisar banco de dados
- Ajuste do banco de dados
- Pré-Processamento
- Análise de frequência
- Seleção de palavra mais importante
- Aplicação de algoritmos
- Avaliação e Visualização



Resultados

Os resultados que obtemos não foram muito positivos
T5 tem uma média de aproximadamente 53%



Proposta

A proposta que seguimos foi a de se fixar na criação das questões, podendo ser abertas ou de múltipla escolha, através de mineração de texto e aprendizagem de máquina



Dataset

Usamos o dataset disponível no site <https://rajpurkar.github.io/SQuAD-explorer/>
Como podemos perceber avaliando o banco ele possui diversos campos mas aqui destacamos os que usaremos:
Question: A pergunta
Context: Um texto que passa um contexto para a pergunta
Answer: A resposta correta

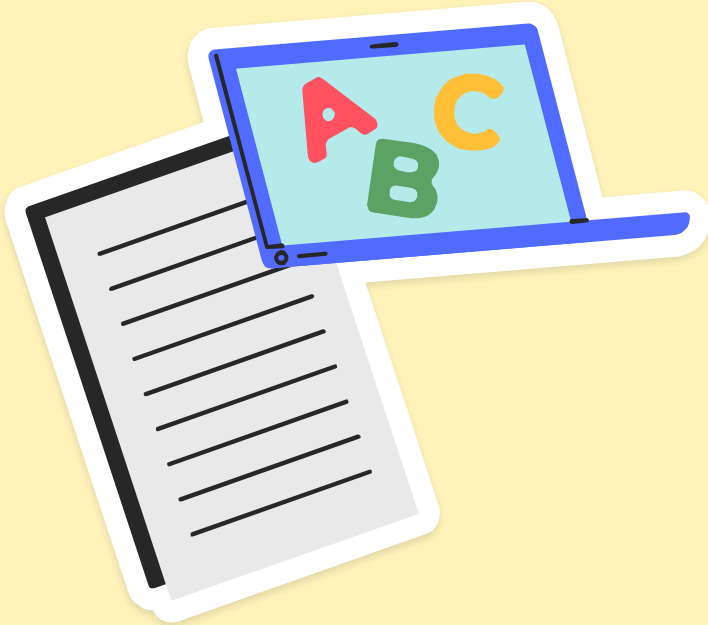


Conclusão

Como indicamos na seção acima os resultados não foram tão positivos, talvez com um processo de seleção de palavra chave diferente traria um melhor resultado



Tipos de Geração de Perguntas



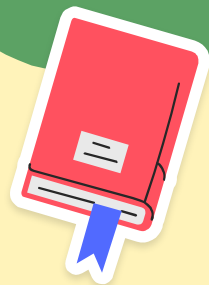
**Categoria -
Nível de
entendimento**

01

Geração baseada na sintaxe.

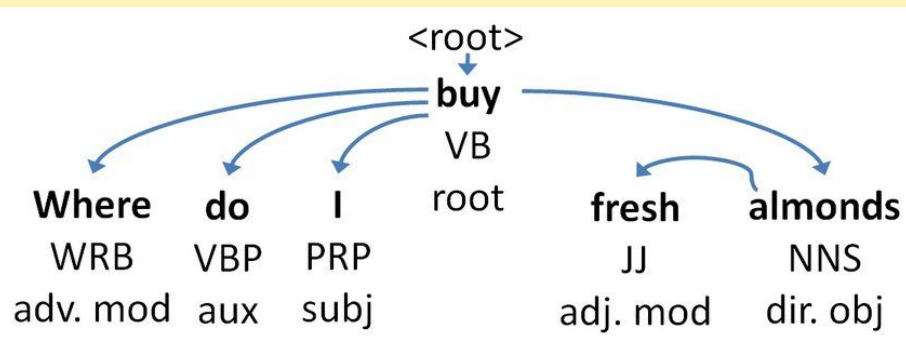


Nível de entendimento (Level of understanding)

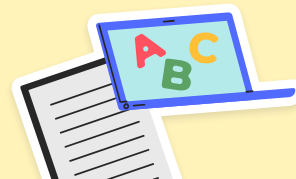


A abordagem **sintática** não precisa entender a **semântica** do input, ou seja ele **não** precisa entender o **sentido** das palavras nem do texto

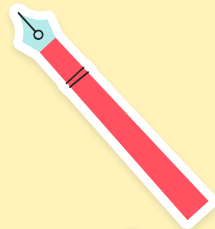
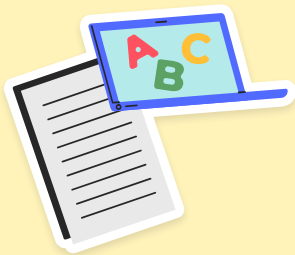
Um exemplo de abordagem **sintática** é a seleção de **respostas erradas** usando o **POS (Part-of-Speech)**



Exemplo de POS



02



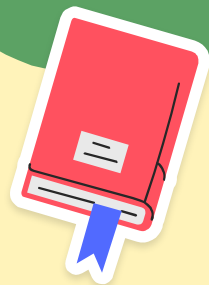
Geração

°baseada na

semântica.

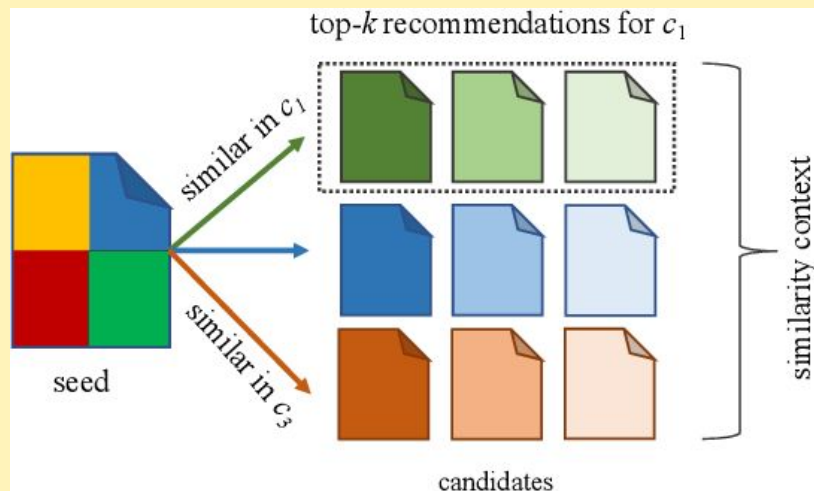


Nível de entendimento (Level of understanding)

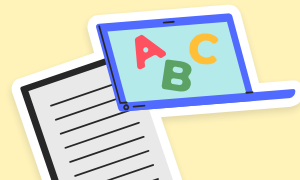


A Abordagem **semântica** precisa de um entendimento mais **aprofundado** do texto. Muitas vezes as **informações** **não estão** de forma **explícita** no **input**

Um exemplo da abordagem **semântica** é o uso de **similaridade contextual** (**contextual similarity**) para a classificação das **respostas erradas**



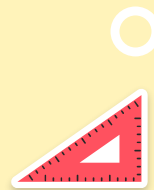
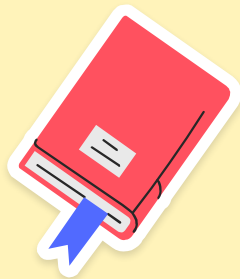
exemplo de contextual similarity





Categoria - Processo de transformação

01



Geração

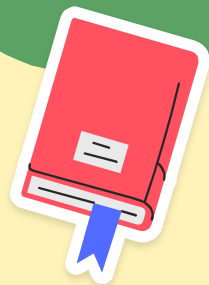
baseada no

template



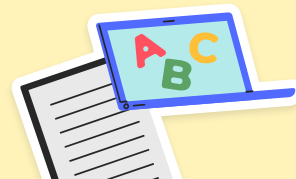


Processo de transformação (Procedure of transformation)

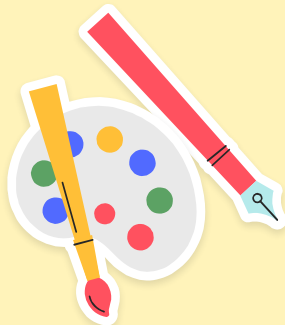


As **questões** são geradas com o uso de **templates**.

Eles definem a **estrutura base** das **questões** usando **texto fixo** e **espaços reservados** que são **substituídos** por inputs para gerar as **questões**



02



Geração

baseada nas

Regras

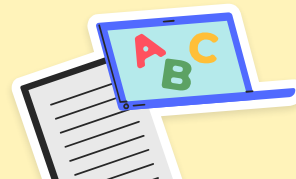
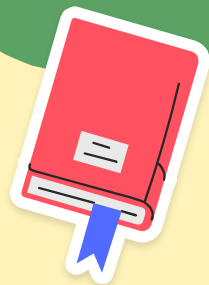
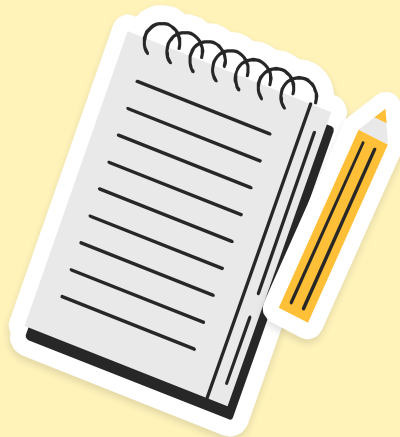




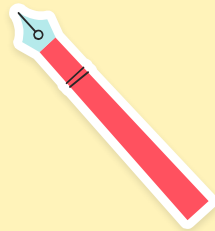
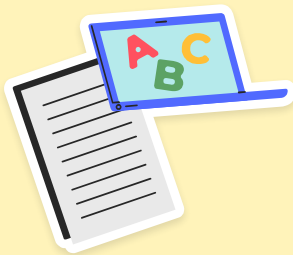
Processo de transformação (Procedure of transformation)

Questões são gerados com o uso de regras

As regras definem como escolher um tipo adequado de questão e como manipular o input para gerar questões



02



Geração baseada na estatística.

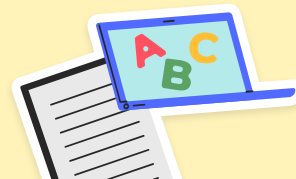
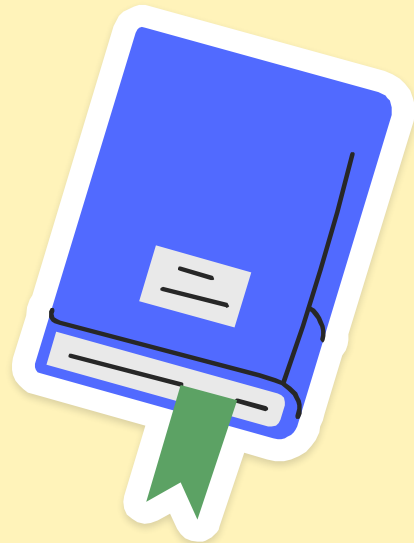




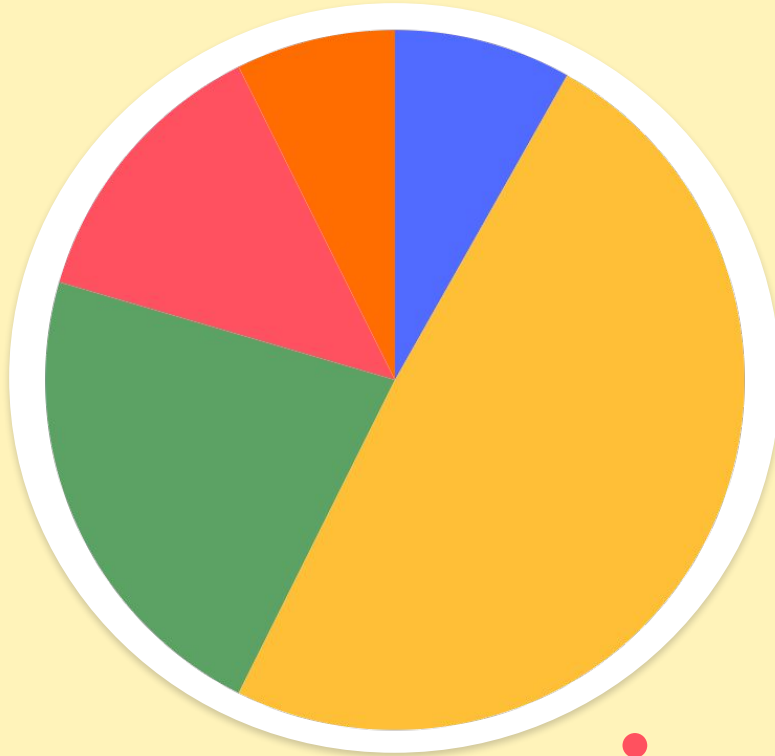
Processo de transformação (Procedure of transformation)

É nesse método que a **geração** de **questões** é aprendida a partir de **dados de treinamento**.

Em um **exemplo** o problema foi encarado como um **sequence-to-sequence prediction problem**



Usabilidade



60 Semântica
10 Sintaxe
27 Template
16 Regras
9 Estatística



Gerando Perguntas

pré-processamento

Básico

Segmentation, Sentence splitting, Tokenization, POS tagging, Coreference resolution e dependendo Entity recognition (NER) e Relation extraction (RE).

Específico

Sentence simplification, Sentence classification, Content selection

Gerando Perguntas

métodos

Stem and correct
answer generation

Feedback
generation

Incorrect options
(i.e. distractor)
generation

Controlling
difficulty

Gerando Perguntas

Stem and correct answer generation

Podem usar
templates, regras,
ou métodos
estatísticos

Transforma frases
afirmativas em
perguntas

Determina o tipo
de questão

Gerando Perguntas

Incorrect options (i.e. distractor) generation

Gera opções de
respostas **erradas**
a questão

A estratégia mais
usada é a baseada
em **similarity**

Outra estratégia
possível é o uso do
word frequency

Gerando Perguntas

Feedback generation

Explica o **porquê** a resposta da questão está certa ou **errada**

Apresentar **materiais de apoio**

Normalmente esse método é **deixado de lado** nos artigos (apenas **um artigo** mencionado utilizou esse método)

Gerando Perguntas

Controlling difficulty

Determina o quão **difícil** ou simples a **questão** está

Uma forma de determinar **difículdade** é através do **percentual de corretude** (percentage correct) que mede quantas pessoas responderam **corretamente** a **questão**

Nossos Resultados Iniciais

Frequência de Palavras

Resultado **positivo**
Média de
aproximadamente
0.866

Similaridade

Resultado **negativo**
Não conseguimos
observar um padrão



De olho no código



Palavra chave

Usamos a **frequência da aparição** das palavras no texto de apoio para escolher a palavra mais importante ou a **palavra chave** do texto

Retirando Ambiguidade



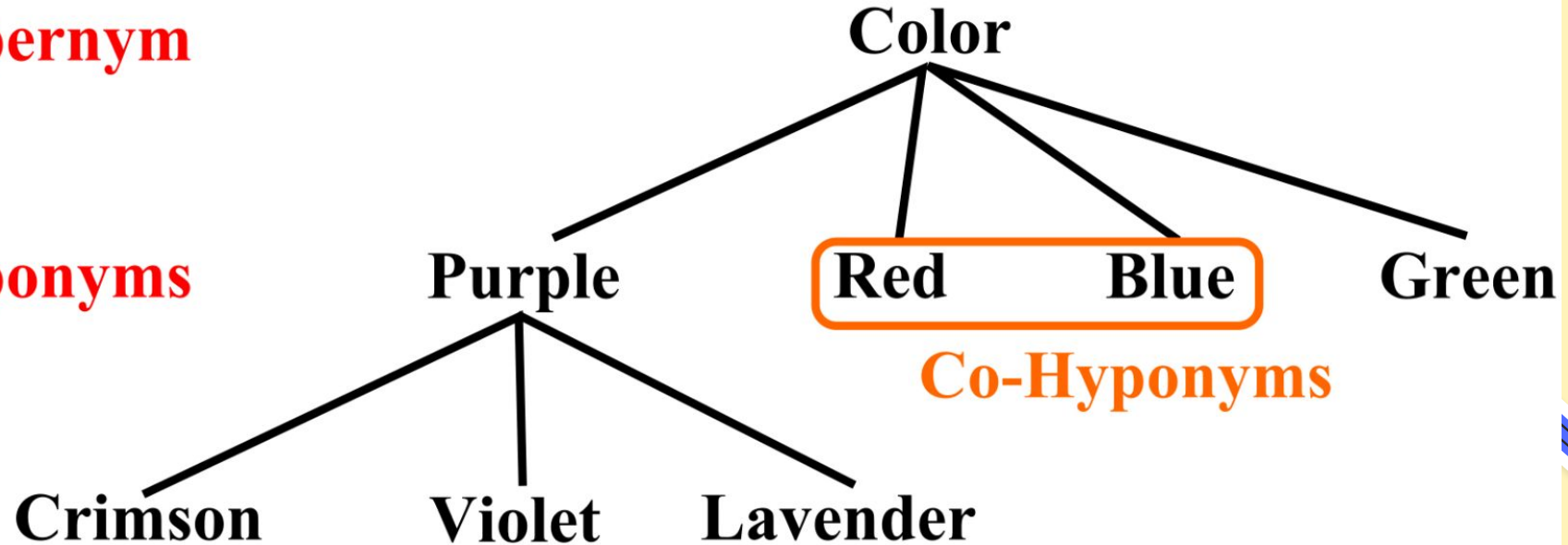
<https://github.com/BPYap/BERT-WSD>



Gerando respostas erradas ○

Hypernym

Hyponyms



BART comparação

	SQuAD 1.1 EM/F1	SQuAD 2.0 EM/F1
BERT	84.1/90.9	79.0/81.8
UniLM	-/-	80.5/83.4
XLNet	89.0 /94.5	86.1/88.8
RoBERTa	88.9/ 94.6	86.5/89.4
BART	88.8/ 94.6	86.1/89.2

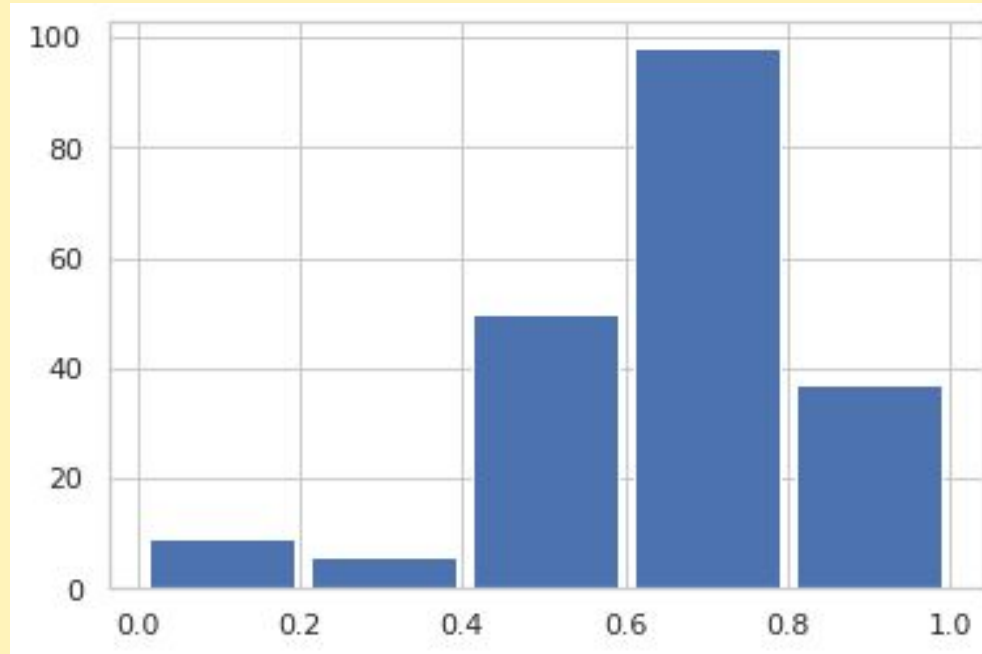
Table 3: BART gives similar results to XLNet and RoBERTa on question answering.

Avaliação

Usamos a **text comparetion** entre o **context** e a **pergunta** gerada para avaliar a geração de perguntas, usando a similaridade do cosseno dos embeddings

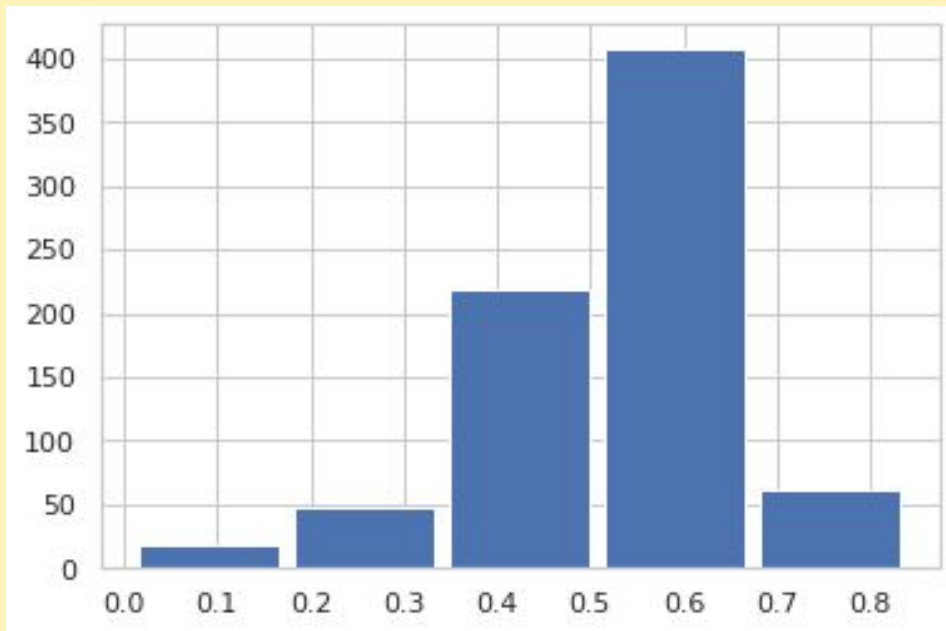
Resultados

T5 (banco na área de ciência) - 66%



Resultados

T5 (área de variedades) - 53%



OBRIGADO!

CRÉDITOS: este modelo de apresentação foi criado pelo **Slidesgo**, e inclui ícones da **Flaticon** e infográficos e imagens da **Freepik**

Por favor, mantenha este slide para atribuição

