



Digital
Transformation
CoLab

The background features several large, abstract geometric shapes in shades of blue and purple. These include a large chevron pointing right, a vertical bar, a large 'X' shape, and various other lines and dots scattered across the dark blue background.

Experiencing
the Future

dtx-colab.pt

Projeto Integrado

Mestrado em Matemática e Computação
2023/2024

Chegar ao DTx

- Campus de Azurém
- Entrar no edifício 01
- Descer escadas encostadas à direita
- Porta de vidro com acesso controlado (ligar para número indicado na porta)
- Virar à esquerda



Questões

- Reuniões de orientação
 - Semanal? Horário?
 - Alternadamente no DTx e via zoom?
- Entrega: 2ª ou 3ª semana de Junho?
- Elementos de avaliação: relatório (~20 pág.) e apresentação?
- On-boarding no DTx: a agendar com os 3 grupos
- Canais de comunicação: slack e e-mail

Análise de sentimento do discurso em transmissões televisivas de jogos de futebol

emanuel.gouveia@dtx-colab.pt

antonio.oliveira@dtx-colab.pt

armindo.lobo@dtx-colab.pt

Motivação

- Em transmissões televisivas em directo, detectar momentos de interesse automaticamente acrescenta valor à transmissão
 - p. ex., permitir "puxar atrás" automaticamente para rever pontos altos
- Para a detecção automática, a faixa de áudio é complementar à imagem, e é rica em características: tom, volume, rapidez, **conteúdo do discurso**
- Eventos desportivos, nomeadamente futebol, geram as maiores audiências para transmissões ao vivo

Dataset

SoccerNet

- Vídeos de transmissões TV de futebol
- Na maior parte dos casos, existe faixa áudio com comentário
- Várias línguas
- 550 jogos de 6 ligas europeias
- Anotados com momentos de destaque
- <https://www.soccer-net.org/home>
- Extrair faixas áudio com moviepy (<https://pypi.org/project/moviepy/>) ou ffmpeg
- Podemos ajudar se tiverem falta de armazenamento ou dificuldade de acesso aos dados

Reconhecimento de discurso

Extrair texto com comentário a partir das faixas áudio

Whisper, da OpenAI

<https://openai.com/research/whisper>

Disponível no huggingface

<https://huggingface.co/openai/whisper-large-v3>

Se encontrarem problemas de memória ou muita lentidão, há versões mais pequenas e rápidas

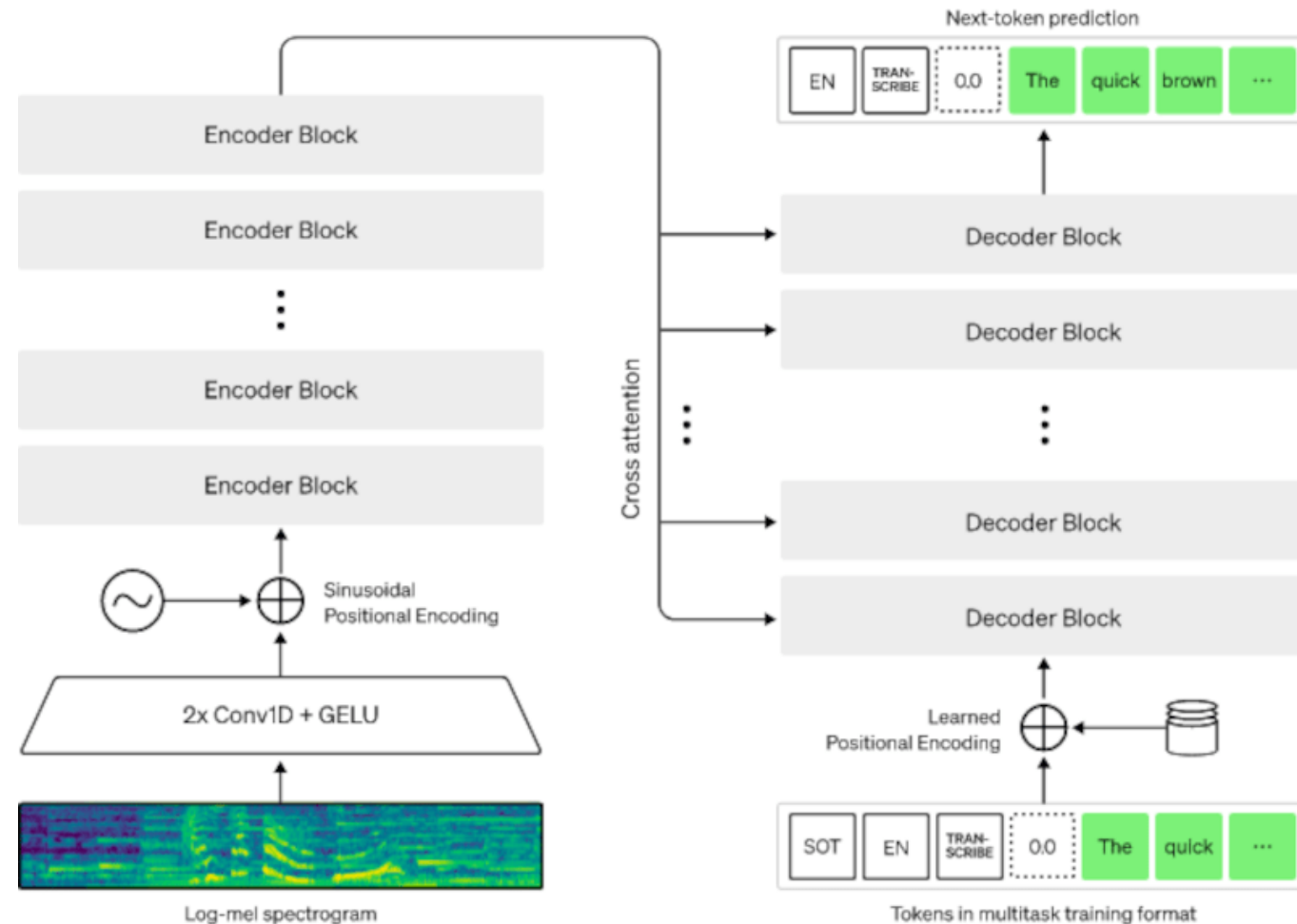
Sugestão:

usem notebooks no Google Colab

<https://colab.research.google.com/>

Como usar modelos do huggingface:

https://huggingface.co/docs/transformers/main_classes/pipelines



Classificação de texto *zero-shot*

Classificar excertos de texto em classes não utilizadas no treino

⚡ Zero-Shot Classification demo

using [facebook/bart-large-mnli](#)

⚙️ Zero-Shot Classification

há uma recuperação de bola, vão sair para o contra-ataque, mas perdem a bola novamente

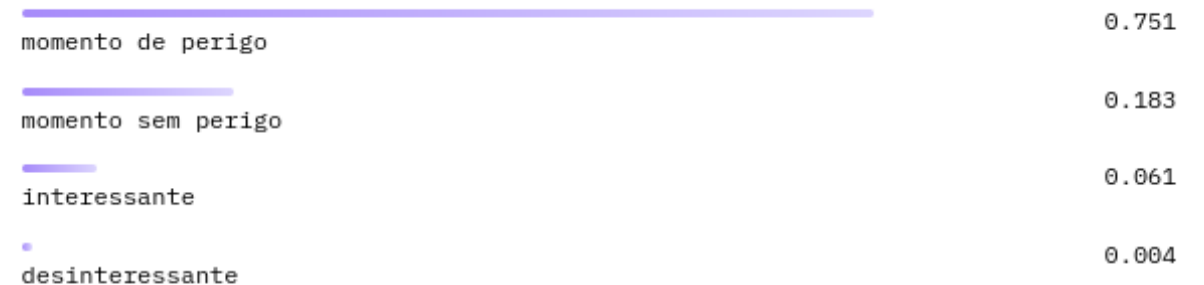
Possible class names (comma-separated)

momento de perigo, momento sem perigo, interessante, desinteressante

☐ Allow multiple true classes

Compute

Computation time on cpu: 1.126 s



</> JSON Output

🔍 Maximize

Classificação de texto *zero-shot*

⚡ Zero-Shot Classification demo

using [facebook/bart-large-mnli](#)

⚙️ Zero-Shot Classification

os jogadores vão passando a bola cá atrás, falta intenção a esta equipa

Possible class names (comma-separated)

momento de perigo, momento sem perigo, interessante, desinteressante

☐ Allow multiple true classes

Compute

Computation time on cpu: 1.134 s



</> JSON Output

🔍 Maximize

⚡ Zero-Shot Classification demo

using [facebook/bart-large-mnli](#)

⚙️ Zero-Shot Classification

esta é uma falta claríssima, o árbitro não assinala, e vemos grandes protestos do público

Possible class names (comma-separated)

momento de perigo, momento sem perigo, interessante, desinteressante

☐ Allow multiple true classes

Compute

Computation time on cpu: 1.235 s



</> JSON Output

🔍 Maximize

Classificação de texto

Modelos de linguagem como classificadores *zero-shot*

p. ex. <https://huggingface.co/google-bert/bert-base-uncased>

Fine-tuning de modelos de linguagem para classificação *zero-shot*

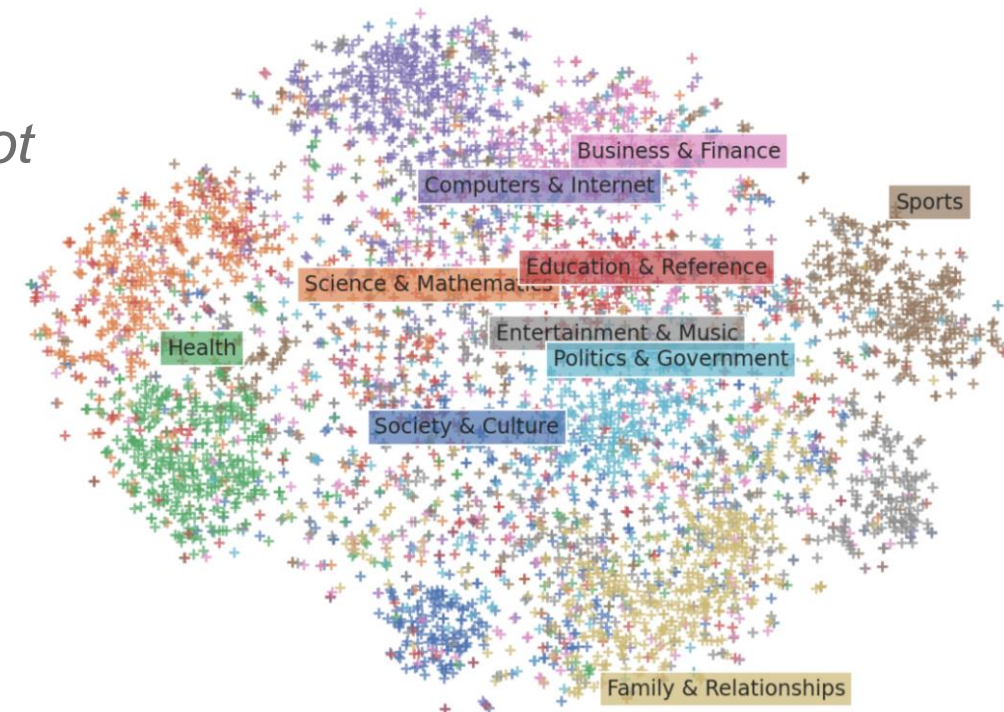
p. ex. <https://huggingface.co/facebook/bart-large-mnli>

Fine-tuning de modelos de linguagem para classificação tradicional (classes conhecidas)

p. ex. <https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student>

Referência genérica

<https://joeddav.github.io/blog/2020/05/29/ZSL.html>



t-SNE visualization of embeddings with SBERT to Wordvec projection. This extra projection step results in

Premise	Label	Hypothesis
A man inspects the uniform of a figure in some East Asian country.	contradiction	The man is sleeping.
An older and younger man smiling.	neutral	Two men are smiling and laughing at the cats playing on the floor.
A soccer game with multiple males playing.	entailment	Some men are playing a sport.

Objectivos

- Utilizar modelos de *deep learning* pré-treinados numa *pipeline* de reconhecimento de discurso
- Treinar e utilizar modelos de *deep learning* para classificação de texto para detectar momentos de destaque em transmissões TV de jogos de futebol
- Comparar três abordagens:
 - Aplicar classificador *zero-shot* pré-treinado
 - Fazer *fine-tuning* para classificação *zero-shot* no *dataset* alvo
 - Fazer *fine-tuning* para classificação tradicional no *dataset* alvo

