

Actividad 3 - BertForQuestionAnswering

November 22, 2023

1 Actividad 3: Entrenamiento de un transformer para Q&A

Alfonso Pineda | A01660394
Mariana Rincón | A01654973
Karla González | A01541526
Salvador Mendoza | A01067783
Álvaro Morán Errejón | A01638034

Fecha de entrega: 27 de Noviembre de 2023

1.1 Instrucciones

- Realice un reporte en equipos de reto respondiendo las siguientes preguntas.
- Obtenga una base de conocimiento o Corpus con información técnica referente a su reto.
- Utilizando un modelo previamente entrenado de “BertForQuestionAnswering”, administre como corpus el texto usado en las actividades anteriores o algún otro diferente.
- Plantee 10 preguntas que el transformer debería de responder con respecto al corpus.
- Obtenga las respuestas de esas 10 preguntas en español e inglés (recuerden que sólo se entrena una vez, la idea es ver las diferentes respuestas con entradas de diferentes idiomas):
 - ¿Hubo alguna diferencia?
 - ¿Qué lenguaje conviene más y por qué?
 - ¿Cuál era el tamaño del corpus?
 - ¿Cuántas respuestas tienen coherencia?
 - ¿Si cambia el corpus y pregunta lo mismo recibirá una respuesta? Demuestre
 - ¿Cuántos lenguajes puede manejar el BERT para resolver preguntas?

1.2 Solución

1.2.1 Importación de datos

Como primer paso, es necesario importar los datos que se van a utilizar para el entrenamiento del modelo. En este caso, se utilizará un corpus de un documento en formato PDF que contienen información sobre la detección anatómica usando landmarks. Para esto, se creó una función que permite leer los documentos PDF y extraer el contenido de los mismos.

El paper que se utilizó para este ejercicio se encuentra en el siguiente enlace: [Database guided detection of anatomical landmark points in 3D images of the heart](#)

```
[14]: import PyPDF2

def read_pdf(file_path):
    with open(file_path, 'rb') as file:
        reader = PyPDF2.PdfReader(file)
        content = ""
        for page in reader.pages:
            content += page.extract_text()
    return content
```

Asimismo, es necesario realizar una función que nos permita realizar un conteo de palabras en el corpus, para poder tener una idea de la cantidad de palabras que se utilizarán para el entrenamiento del modelo y la frecuencia de las mismas.

```
[15]: from collections import Counter
import re

def contar_palabras(texto):
    palabras = re.findall(r'\w+', texto.lower())
    contador = Counter(palabras)
    return contador
```

Utilizamos las dos funciones anteriores para poder obtener el corpus y el conteo de palabras del mismo.

```
[16]: file_content = read_pdf("datasets/KaravidesSPIE-MI2010Approved.pdf")
word_counts = contar_palabras(file_content)
```

Obtenemos el tamaño total del corpus, el cual es de 4,168 palabras y la cantidad de palabras únicas, que es de 1,072.

```
[17]: print("Cantidad de palabras distintas: ", len(word_counts))
print("Cantidad de palabras totales: ", sum(word_counts.values()))
```

Cantidad de palabras distintas: 1072

Cantidad de palabras totales: 4168

Posteriormente, obtenemos las 10 palabras más frecuentes del corpus.

```
[18]: # 10 palabras mas usadas en el pdf
print(word_counts.most_common(10))
```

```
[('the', 229), ('and', 119), ('of', 106), ('to', 80), ('in', 68), ('a', 67),
('2', 48), ('are', 46), ('3', 42), ('all', 41)]
```

1.2.2 Definición y entrenamiento del modelo

Para el entrenamiento del modelo, se utilizó el modelo pre-entrenado de BertForQuestionAnswering, el cual se encuentra en la librería de transformers. Este modelo se entrenó con el corpus que se obtuvo en el paso anterior.

A continuación, se muestra el código utilizado para la carga del modelo y su tokenizador correspondiente.

```
[19]: from transformers import BertForQuestionAnswering, BertTokenizer

model_name = "bert-large-uncased-whole-word-masking-finetuned-squad"
model = BertForQuestionAnswering.from_pretrained(model_name)
tokenizer = BertTokenizer.from_pretrained(model_name)
```

Some weights of the model checkpoint at bert-large-uncased-whole-word-masking-finetuned-squad were not used when initializing BertForQuestionAnswering:

```
['bert.pooler.dense.bias', 'bert.pooler.dense.weight']
```

- This IS expected if you are initializing BertForQuestionAnswering from the checkpoint of a model trained on another task or with another architecture (e.g. initializing a BertForSequenceClassification model from a BertForPreTraining model).

- This IS NOT expected if you are initializing BertForQuestionAnswering from the checkpoint of a model that you expect to be exactly identical (initializing a BertForSequenceClassification model from a BertForSequenceClassification model).

A continuación, como primer paso se realiza la definición de las preguntas que se le harán al modelo. Para esto, se creó una lista con las preguntas relacionadas al corpus ingresado. Estas preguntas servirán como consultas para obtener información específica del contexto.

Posteriormente se hace la segmentación del corpus, dado que el modelo tiene limitaciones en la cantidad de tokens que puede procesar, se divide el contexto en segmentos manejables. Este enfoque asegura que se abarque la totalidad del contenido, evitando la pérdida de información relevante. Para esto, se creó una función que permite dividir el corpus en segmentos de 512 tokens, que es el máximo que puede procesar el modelo.

Una vez definidas las preguntas y segmentado el corpus, se itera a través de cada pregunta y, para cada pregunta, se subdividen los segmentos del contexto. Esto facilita la adaptación del modelo a porciones más pequeñas del texto original.

Para cada combinación de pregunta y segmento, se preparan los datos de entrada para el modelo mediante la tokenización. Se utiliza la tokenización de pregunta y segmento, y se obtienen las puntuaciones de inicio y fin para determinar la posición de la respuesta en el segmento.

Las puntuaciones de inicio y fin se utilizan para identificar la posición de inicio y fin de la respuesta dentro del segmento. Si se detecta una respuesta válida, se convierten los tokens de la respuesta de nuevo a texto y se imprime el resultado.

Finalmente, se obtiene la respuesta a la pregunta en cada segmento y se almacena en una lista. Esta lista contiene las respuestas a las 10 preguntas realizadas.

```
[20]: import torch

questions = [
    "How can automated landmark detection be invaluable in the analysis of_
    ↪real-time three-dimensional echocardiograms?",
```

```

    "Why is the standardization of views critical in the interpretation of
    ↪echocardiographic images, and how does variability in the choice of
    ↪anatomical views affect diagnosis?",
    "How do manual and automated methods compare in detecting anatomical
    ↪landmark points in echocardiography?",
    "How does the classification method work for detecting landmark points in
    ↪heart images, and what techniques are used?",
    "What were the results of automated landmark detection, and how do these
    ↪results compare with other existing methods?",
    "How does variability in echocardiographic images affect the accuracy of
    ↪diagnosis?",
    "What are the main challenges in standardizing anatomical views in
    ↪echocardiography?",
    "What are the latest technological advances in three-dimensional cardiac
    ↪ultrasound?",
    "What has been the impact of introducing automated methods in
    ↪echocardiography practice?",
    "How do manual and automated methods compare in terms of accuracy and
    ↪reliability in detecting anatomical landmark points in echocardiography?"
]

context = file_content

for question in questions:
    max_length = 512 - len(tokenizer.encode(question, add_special_tokens=True))
    ↪- 1
    context_segments = []

    start = 0
    end = max_length
    while end < len(context):
        while end < len(context) and context[end] != ' ':
            end += 1
        context_segments.append(context[start:end])
        start = end
        end = start + max_length
    context_segments.append(context[start:])

    for segment in context_segments:
        inputs = tokenizer.encode_plus(question, segment,
        ↪add_special_tokens=True, return_tensors='pt')
        input_ids = inputs['input_ids'][0]

        with torch.no_grad():
            outputs = model(**inputs)
            answer_start_scores = outputs.start_logits

```

```

answer_end_scores = outputs.end_logits

answer_start = torch.argmax(answer_start_scores)
answer_end = torch.argmax(answer_end_scores) + 1

if answer_start < answer_end:
    answer = tokenizer.convert_tokens_to_string(tokenizer.
↪convert_ids_to_tokens(input_ids[answer_start:answer_end]))
    print(f"Pregunta: {question}\nRespuesta: {answer}\n")
    break

```

Pregunta: How can automated landmark detection be invaluable in the analysis of real-time three-dimensional echocardiograms?

Respuesta: database guided detection of anatomical landmark points in 3d images of the heart

Pregunta: Why is the standardization of views critical in the interpretation of echocardiographic images, and how does variability in the choice of anatomical views affect diagnosis?

Respuesta: variability in the choice of anatomical views affect diagnosis

Pregunta: How do manual and automated methods compare in detecting anatomical landmark points in echocardiography?

Respuesta: manual and automated methods compare in detecting anatomical landmark points in echocardiography ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided detection

Pregunta: How does the classification method work for detecting landmark points in heart images, and what techniques are used?

Respuesta: database guided detection

Pregunta: What were the results of automated landmark detection, and how do these results compare with other existing methods?

Respuesta: compare with other existing methods

Pregunta: How does variability in echocardiographic images affect the accuracy of diagnosis?

Respuesta: variability in echocardiographic images affect the accuracy of diagnosis ? [SEP]

Pregunta: What are the main challenges in standardizing anatomical views in echocardiography?

Respuesta: database guided detection of anatomical landmark points

Pregunta: What are the latest technological advances in three-dimensional cardiac ultrasound?

Respuesta: guided detection of anatomical landmark points

Pregunta: What has been the impact of introducing automated methods in echocardiography practice?

Respuesta: database guided detection of anatomical landmark points in 3d images of the heart

Pregunta: How do manual and automated methods compare in terms of accuracy and reliability in detecting anatomical landmark points in echocardiography?

Respuesta: accuracy and reliability

Debido a que las respuestas anteriormente obtenidas no fueron muy precisas, se decidió ocupar un nuevo enfoque que busca mejorar la precisión de las respuestas mediante una estrategia de segmentación de contexto más refinada. Se ha implementado una función denominada `split_into_segments` para dividir el texto en segmentos más pequeños, abordando las limitaciones previas en la precisión de las respuestas.

La función `split_into_segments` toma el texto del contexto y lo divide en segmentos más pequeños de manera más precisa. En este caso, se ha implementado un enfoque basado en palabras, asegurando que cada segmento se ajuste a una longitud máxima definida (250 en este caso). Este método permite una segmentación más coherente y centrada en el contenido textual.

Se mantiene el proceso de iterar sobre cada pregunta y, para cada pregunta, se itera sobre los nuevos segmentos generados. La tokenización y el procesamiento de preguntas y segmentos siguen siendo fundamentales, pero ahora se realiza en función de la segmentación mejorada proporcionada por la función `split_into_segments`.

Con la nueva segmentación, se espera obtener respuestas más precisas. Se sigue utilizando la lógica de identificación de la posición de inicio y fin de la respuesta en cada segmento, y se imprime el resultado si se encuentra una respuesta válida.

```
[21]: questions = [
    "How can automated landmark detection be invaluable in the analysis of_
    ↪real-time three-dimensional echocardiograms?",
    "Why is the standardization of views critical in the interpretation of_
    ↪echocardiographic images, and how does variability in the choice of_
    ↪anatomical views affect diagnosis?",
    "How do manual and automated methods compare in detecting anatomical_
    ↪landmark points in echocardiography?",
    "How does the classification method work for detecting landmark points in_
    ↪heart images, and what techniques are used?",
    "What were the results of automated landmark detection, and how do these_
    ↪results compare with other existing methods?",
    "How does variability in echocardiographic images affect the accuracy of_
    ↪diagnosis?",
    "What are the main challenges in standardizing anatomical views in_
    ↪echocardiography?",
```

```

    "What are the latest technological advances in three-dimensional cardiac_
    ↪ultrasound?",
    "What has been the impact of introducing automated methods in_
    ↪echocardiography practice?",
    "How do manual and automated methods compare in terms of accuracy and_
    ↪reliability in detecting anatomical landmark points in echocardiography?"
]

context = file_content

def split_into_segments(text, max_length):
    words = text.split()
    segments = []
    current_segment = ""

    for word in words:
        if len(current_segment) + len(word) + 1 <= max_length:
            current_segment += word + " "
        else:
            segments.append(current_segment.strip())
            current_segment = word + " "
    segments.append(current_segment.strip())
    return segments

context_segments = split_into_segments(context, 250) #De igual forma que el_
    ↪metodo anterior dividimos el contexto en segmentos

```

A continuación, se hace uso de las funciones `get_top_answers` y `answer_question` para realizar la tokenización y el procesamiento necesario. A continuación, se explica de manera más detallada el funcionamiento de cada una de estas funciones:

- **get_top_answers** Recibe las posiciones de inicio y fin de respuestas potenciales, así como los identificadores de entrada (`input_ids`). Convierte los identificadores de respuesta en texto y los almacena en una lista.
- **answer_question** Tokeniza la pregunta y el segmento del contexto, preparando la entrada para el modelo. Obtiene las puntuaciones de inicio y fin del modelo para evaluar las posibles ubicaciones de respuesta, selecciona las mejores 'N' posiciones tanto para el inicio como para el fin de la respuesta y utiliza la función `get_top_answers` para convertir estas posiciones en respuestas de texto.

Posteriormente se itera sobre cada pregunta en la lista `questions`, para cada pregunta itera sobre los segmentos del contexto utilizando la función de segmentación mejorada. Posteriormente llama a la función `answer_question` para obtener las tres mejores respuestas para la pregunta actual y el segmento actual y por último almacena dichas respuestas en la lista `all_answers`.

Finalmente, se imprimen las preguntas junto con las tres mejores respuestas obtenidas para cada una.

```

[22]: import numpy as np

def get_top_answers(possible_starts, possible_ends, input_ids):
    answers = []
    for start, end in zip(possible_starts, possible_ends):
        answer = tokenizer.convert_tokens_to_string(tokenizer.
↪convert_ids_to_tokens(input_ids[start:end+1]))
        answers.append(answer)
    return answers

def answer_question(question, context_segment, topN=3):
    inputs = tokenizer.encode_plus(question, context_segment,
↪add_special_tokens=True, return_tensors="pt")
    input_ids = inputs["input_ids"].tolist()[0]

    model_out = model(**inputs)

    answer_start_scores = model_out["start_logits"]
    answer_end_scores = model_out["end_logits"]

    possible_starts = np.argsort(answer_start_scores.cpu().detach().numpy()).
↪flatten()[::-1][:topN]
    possible_ends = np.argsort(answer_end_scores.cpu().detach().numpy()).
↪flatten()[::-1][:topN]

    answers = get_top_answers(possible_starts, possible_ends, input_ids)

    return answers

context = file_content
all_answers = []
for question in questions:
    segment_answers = []
    for segment in context_segments:
        if len(tokenizer.encode(question, segment, add_special_tokens=True)) <=
↪512:
            answers = answer_question(question, segment, topN=3)
            for answer in answers:
                if answer.strip() and answer not in segment_answers:
                    segment_answers.append(answer)
                if len(segment_answers) >= 3:
                    break
            if len(segment_answers) >= 3:
                break

    all_answers.append((question, segment_answers))

```



```

for question, answers in all_answers:
    print(f"Question: {question}")
    for i, answer in enumerate(answers, 1):
        print(f"Answer {i}: {answer}")
    print()

```

Question: How can automated landmark detection be invaluable in the analysis of real-time three-dimensional echocardiograms?

Answer 1: database guided detection of anatomical landmark points in 3d images of the heart

Answer 2: guided detection

Answer 3: imaging

Question: Why is the standardization of views critical in the interpretation of echocardiographic images, and how does variability in the choice of anatomical views affect diagnosis?

Answer 1: variability in the choice of anatomical views affect diagnosis

Answer 2: how does variability in the choice of anatomical views

Answer 3: choice of anatomical views

Question: How do manual and automated methods compare in detecting anatomical landmark points in echocardiography?

Answer 1: guided detection

Answer 2: manual and automated methods compare in detecting anatomical landmark points in echocardiography ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided

Answer 3: 3d images of the heart

Question: How does the classification method work for detecting landmark points in heart images, and what techniques are used?

Answer 1: guided detection

Answer 2: database guided

Answer 3: detection of anatomical landmark points in 3d images of the heart

Question: What were the results of automated landmark detection, and how do these results compare with other existing methods?

Answer 1: compare with other existing methods ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided detection of anatomical landmark points in 3d images of the heart

Answer 2: database guided detection of anatomical landmark points in 3d images

Answer 3: compare with other existing methods

Question: How does variability in echocardiographic images affect the accuracy of diagnosis?

Answer 1: variability in echocardiographic images affect the accuracy of

diagnosis ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided detection

Answer 2: affect the accuracy

Answer 3: variability in echocardiographic images affect the accuracy

Question: What are the main challenges in standardizing anatomical views in echocardiography?

Answer 1: what are the main challenges in standardizing anatomical views in echocardiography ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided detection of anatomical landmark points

Answer 2: database guided detection

Answer 3: guided detection of anatomical landmark points in 3d images of the heart

Question: What are the latest technological advances in three-dimensional cardiac ultrasound?

Answer 1: guided detection of anatomical landmark points

Answer 2: database guided detection of anatomical landmark points in 3d images of the heart

Answer 3: detection

Question: What has been the impact of introducing automated methods in echocardiography practice?

Answer 1: database guided detection of anatomical landmark points in 3d images of the heart

Answer 2: guided detection

Answer 3: what has been the impact of introducing automated methods in echocardiography practice ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided detection of anatomical landmark points

Question: How do manual and automated methods compare in terms of accuracy and reliability in detecting anatomical landmark points in echocardiography?

Answer 1: accuracy and reliability

Answer 2: manual and automated methods compare in terms of accuracy and reliability in detecting anatomical landmark points in echocardiography ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided detection

Answer 3: reliability in detecting anatomical landmark points in echocardiography ? [SEP] see discussions , st ats , and author pr ofiles f or this public ation at : [https : / / www . researchgate . ne t / public ation / 221624938](https://www.researchgate.net/publication/221624938) database guided detection of anatomical landmark points in 3d images of the heart

Posteriormente, realizamos el mismo proceso pero con las preguntas en español y observamos los resultados obtenidos.

```
[23]: questions = [
    "¿Cómo puede ser invaluable la detección automática de puntos de referencia_
    ↪ en el análisis de ecocardiogramas tridimensionales en tiempo real?",
    "¿Por qué es crítica la estandarización de vistas en la interpretación de_
    ↪ imágenes ecocardiográficas, y cómo afecta la variabilidad en la elección de_
    ↪ vistas anatómicas al diagnóstico?",
    "¿Cómo se comparan los métodos manuales y automáticos en la detección de_
    ↪ puntos de referencia anatómicos en la ecocardiografía?",
    "¿Cómo funciona el método de clasificación para detectar puntos de_
    ↪ referencia en imágenes cardíacas y qué técnicas se utilizan?",
    "¿Cuáles fueron los resultados de la detección automática de puntos de_
    ↪ referencia y cómo se comparan estos resultados con otros métodos existentes?
    ↪",
    "¿Cómo afecta la variabilidad en las imágenes ecocardiográficas a la_
    ↪ precisión del diagnóstico?",
    "¿Cuáles son los principales desafíos en la estandarización de vistas_
    ↪ anatómicas en la ecocardiografía?",
    "¿Cuáles son los últimos avances tecnológicos en el ultrasonido cardíaco_
    ↪ tridimensional?",
    "¿Cuál ha sido el impacto de la introducción de métodos automatizados en la_
    ↪ práctica de la ecocardiografía?",
    "¿Cómo se comparan los métodos manuales y automáticos en términos de_
    ↪ precisión y fiabilidad en la detección de puntos de referencia anatómicos en_
    ↪ la ecocardiografía?"
]

context = file_content
all_answers = []
for question in questions:
    segment_answers = []
    for segment in context_segments:
        if len(tokenizer.encode(question, segment, add_special_tokens=True)) <= 512:
            answers = answer_question(question, segment, topN=3)
            for answer in answers:
                if answer.strip() and answer not in segment_answers:
                    segment_answers.append(answer)
                if len(segment_answers) >= 3:
                    break
            if len(segment_answers) >= 3:
                break
    all_answers.append((question, segment_answers))
```

```

for question, answers in all_answers:
    print(f"Question: {question}")
    for i, answer in enumerate(answers, 1):
        print(f"Answer {i}: {answer}")
    print()

```

Question: ¿Cómo puede ser invaluable la detección automática de puntos de referencia en el análisis de ecocardiogramas tridimensionales en tiempo real?

Answer 1: database guided detection of anatomical landmark points in 3d images of the heart

Answer 2: [SEP] international symposium on biomedical imaging : from nano to macro . ieee international symposium on biomedical imaging · march 2010 doi : 10 . 1109 / isbi . 2010 . 5490182 · source : dblp citations 12reads 271 5 authors , including : pavel paclik pr [SEP]

Answer 3: [CLS] ¿ como puede ser invaluable la deteccion automatica de puntos de referencia en el analisis de ecocardiogramas tridimensionales en tiempo real ? [SEP]

Question: ¿Por qué es crítica la estandarización de vistas en la interpretación de imágenes ecocardiográficas, y cómo afecta la variabilidad en la elección de vistas anatómicas al diagnóstico?

Answer 1: por que es critica la estandarizacion de vistas en la interpretacion de imagenes ecocardiograficas , y como afecta la variabilidad en la eleccion de vistas anatomicas al diagnostico ? [SEP]

Answer 2: [SEP]

Answer 3: como afecta la variabilidad en la eleccion de vistas anatomicas al diagnostico ? [SEP]

Question: ¿Cómo se comparan los métodos manuales y automáticos en la detección de puntos de referencia anatómicos en la ecocardiografía?

Answer 1: database guided detection of anatomical landmark points in 3d images of the heart

Answer 2: [SEP] international symposium on biomedical imaging : from nano to macro . ieee international symposium on biomedical imaging · march 2010 doi : 10 . 1109 / isbi . 2010 . 5490182 · source : dblp citations 12reads 271 5 authors , including : pavel paclik pr [SEP]

Answer 3: [SEP]

Question: ¿Cómo funciona el método de clasificación para detectar puntos de referencia en imágenes cardíacas y qué técnicas se utilizan?

Answer 1: paper in proceedings / ieee

Answer 2: database guided detection of anatomical landmark points in 3d images of the heart conference paper

Answer 3: 221624938

Question: ¿Cuáles fueron los resultados de la detección automática de puntos de

referencia y cómo se comparan estos resultados con otros métodos existentes?

Answer 1: guided detection of anatomical landmark points in 3d images of the heart

Answer 2: estos resultados con otros metodos existentes

Answer 3: como se comparan estos resultados con otros metodos existentes ? [SEP]

Question: ¿Cómo afecta la variabilidad en las imágenes ecocardiográficas a la precisión del diagnóstico?

Answer 1: guided detection

Answer 2: database guided detection of anatomical landmark points

Answer 3: 2010

Question: ¿Cuáles son los principales desafíos en la estandarización de vistas anatómicas en la ecocardiografía?

Answer 1: guided detection of anatomical landmark points in 3d images of the heart

Answer 2: database guided detection of anatomical landmark points

Answer 3: [SEP] international symposium on biomedical imaging : from nano to macro . ieee international symposium on biomedical imaging · march 2010 doi : 10 . 1109 / isbi . 2010 . 5490182 · source : dblp citations 12reads 271 5 author s , including : pavel paclik pr [SEP]

Question: ¿Cuáles son los últimos avances tecnológicos en el ultrasonido cardíaco tridimensional?

Answer 1: guided detection of anatomical landmark points

Answer 2: database guided detection

Answer 3: detection of anatomical landmark points in 3d images of the heart

Question: ¿Cuál ha sido el impacto de la introducción de métodos automatizados en la práctica de la ecocardiografía?

Answer 1: discussions , stats , and author profiles for this publication at : https://www.researchgate.net/publication/221624938_database_guided_detection_of_anatomical_landmark_points_in_3d_images_of_the_heart_conference_paper_in_proceedings_ieee [SEP]

Answer 2: 2010

Answer 3: [SEP]

Question: ¿Cómo se comparan los métodos manuales y automáticos en términos de precisión y fiabilidad en la detección de puntos de referencia anatómicos en la ecocardiografía?

Answer 1: [SEP] see discussions , stats , and author profiles for this publication at : https://www.researchgate.net/publication/221624938_database_guided_detection_of_anatomical_landmark_points_in_3d_images_of_the_heart_conference_paper_in_proceedings_ieee [SEP]

Answer 2: [SEP] international symposium on biomedical imaging : from nano to macro . ieee international symposium on biomedical imaging · march 2010 doi : 10 . 1109 / isbi . 2010 . 5490182 · source : dblp citations 12reads 271 5 author s , including : pavel paclik pr [SEP]

Answer 3: [SEP]

Por último, entrenamos el modelo con un corpus distinto y realizamos las mismas preguntas en inglés para observar los resultados obtenidos. El nuevo corpus utilizado se encuentra en el siguiente enlace: [Reciprocal Landmark Detection and Tracking with Extremely Few Annotations](#)

```
[24]: file_content = read_pdf("datasets/
↳Lin_Reciprocal_Landmark_Detection_and_Tracking_With_Extremely_Few_Annotations_CVPR_2021_paper.pdf")
context = file_content
questions = [
    "How can automated landmark detection be invaluable in the analysis of
↳real-time three-dimensional echocardiograms?",
    "Why is the standardization of views critical in the interpretation of
↳echocardiographic images, and how does variability in the choice of
↳anatomical views affect diagnosis?",
    "How do manual and automated methods compare in detecting anatomical
↳landmark points in echocardiography?",
    "How does the classification method work for detecting landmark points in
↳heart images, and what techniques are used?",
    "What were the results of automated landmark detection, and how do these
↳results compare with other existing methods?",
    "How does variability in echocardiographic images affect the accuracy of
↳diagnosis?",
    "What are the main challenges in standardizing anatomical views in
↳echocardiography?",
    "What are the latest technological advances in three-dimensional cardiac
↳ultrasound?",
    "What has been the impact of introducing automated methods in
↳echocardiography practice?",
    "How do manual and automated methods compare in terms of accuracy and
↳reliability in detecting anatomical landmark points in echocardiography?"
]
def split_into_segments(text, max_length):
    words = text.split()
    segments = []
    current_segment = ""

    for word in words:
        if len(current_segment) + len(word) + 1 <= max_length:
            current_segment += word + " "
        else:
            segments.append(current_segment.strip())
            current_segment = word + " "
    segments.append(current_segment.strip())
    return segments
```

```

context_segments = split_into_segments(context, 250) #De igual forma que el
↪ metodo anterior dividimos el contexto en segmentos

all_answers = []
for question in questions:
    segment_answers = []
    for segment in context_segments:
        if len(tokenizer.encode(question, segment, add_special_tokens=True)) <=
        ↪512:
            answers = answer_question(question, segment, topN=3)
            for answer in answers:
                if answer.strip() and answer not in segment_answers:
                    segment_answers.append(answer)
                if len(segment_answers) >= 3:
                    break
            if len(segment_answers) >= 3:
                break

    all_answers.append((question, segment_answers))

for question, answers in all_answers:
    print(f"Question: {question}")
    for i, answer in enumerate(answers, 1):
        print(f"Answer {i}: {answer}")
    print()

```

Question: How can automated landmark detection be invaluable in the analysis of real-time three-dimensional echocardiograms?

Answer 1: extremely few annotations

Answer 2: real - time three - dimensional echocardiograms ? [SEP]

Answer 3: analysis of real - time three - dimensional echocardiograms ? [SEP]
the university of british columbia , canada 2vancouver general hospital , canada
3department of medicine , the university of british columbia 4the division of
cardiology , the university of british columbia { jianzhelin , ghazal , fatemeht
, mohammadj , [SEP]

Question: Why is the standardization of views critical in the interpretation of echocardiographic images, and how does variability in the choice of anatomical views affect diagnosis?

Answer 1: variability in the choice of anatomical views affect diagnosis

Answer 2: reciprocal landmark detection and tracking with extremely few annotations

Answer 3: choice of anatomical views

Question: How do manual and automated methods compare in detecting anatomical landmark points in echocardiography?

Answer 1: extremely few annotations

Answer 2: few

Answer 3: reciprocal

Question: How does the classification method work for detecting landmark points in heart images, and what techniques are used?

Answer 1: reciprocal landmark detection and tracking with extremely few annotations

Answer 2: classification

Answer 3: what techniques are used ? [SEP]

Question: What were the results of automated landmark detection, and how do these results compare with other existing methods?

Answer 1: extremely few annotations

Answer 2: few

Answer 3: compare with other existing methods

Question: How does variability in echocardiographic images affect the accuracy of diagnosis?

Answer 1: variability in echocardiographic images affect the accuracy of diagnosis ? [SEP] reciprocal landmark detection and tracking with extremely few annotations

Answer 2: accuracy

Answer 3: variability in echocardiographic images affect the accuracy

Question: What are the main challenges in standardizing anatomical views in echocardiography?

Answer 1: reciprocal landmark detection and tracking with extremely few annotations

Answer 2: landmark detection

Answer 3: what are the main challenges in standardizing anatomical views in echocardiography ? [SEP]

Question: What are the latest technological advances in three-dimensional cardiac ultrasound?

Answer 1: reciprocal landmark detection and tracking with extremely few annotations

Answer 2: landmark detection and tracking

Answer 3: tracking with extremely few annotations jianzhe lin1ghazal sahebzamani1christina luong2 , 3

Question: What has been the impact of introducing automated methods in echocardiography practice?

Answer 1: reciprocal landmark detection and tracking with extremely few annotations

Answer 2: what has been the impact of introducing automated methods in echocardiography practice ? [SEP]

Answer 3: canada 2vancouver general hospital , canada 3department of medicine ,

the university of british columbia 4the division of cardiology , the university of british columbia { jianzhelin , ghazal , fatemeht , mohammadj , [SEP]

Question: How do manual and automated methods compare in terms of accuracy and reliability in detecting anatomical landmark points in echocardiography?

Answer 1: extremely few annotations

Answer 2: accuracy and reliability

Answer 3: manual and automated methods compare in terms of accuracy and reliability in detecting anatomical landmark points in echocardiography ? [SEP] reciprocal landmark detection and tracking with extremely few

2 Conclusiones

1. ¿Hubo alguna diferencia?

Hubo una clara diferencia entre los dos porque el corpus que utilizamos es en inglés lo cual hace que tenga mejor rendimiento el modelo con las preguntas en inglés. Esto puede deberse a diferencias en la calidad del modelo pre-entrenado en cada idioma y la forma en la que se procesan las preguntas y el contexto en cada idioma. Adicionalmente, en algunas preguntas, las respuestas en español incluyen tokens adicionales como “[CLS]” o “[SEP]”.

2. ¿Qué lenguaje conviene más y por qué?

Inglés, porque el corpus utilizado es un texto en ese idioma, al igual que el modelo pre-entrenado utilizado “BertForQuestionAnswering”, por lo que las preguntas en inglés pueden tener respuestas más coherentes y precisas. En cambio, si el corpus y el modelo fueran específicos para el español, podría ser más apropiado realizar las preguntas en español.

3. ¿Cuál era el tamaño del corpus?

Es un documento PDF de 8 páginas con 1072 palabras únicas, dando un total de 4168 palabras totales.

4. ¿Cuántas respuestas tienen coherencia?

La mayoría de las preguntas tienen un tipo de respuesta coherente sin embargo son respuestas muy cortas que no expanden la idea principal.

5. ¿Si cambia el corpus y pregunta lo mismo recibirá una respuesta? Demuestre

Al emplear un corpus distinto y plantear las mismas preguntas, se evidencia un cambio en las respuestas obtenidas. Esta variación se atribuye a que el nuevo corpus aborda temáticas diferentes en comparación con el corpus anterior. Mientras que el primero se centra en la conceptualización de landmarks, el segundo se enfoca en el proceso de desarrollo de un modelo UNet destinado a la detección de dichos landmarks.

6. ¿Cuántos lenguajes puede manejar el BERT para resolver preguntas?

De acuerdo con el repositorio en GitHub “Extending-Google-BERT-as-Question-and-Answering-model-and-Chatbot”, el modelo preentrenado puede manejar 104 lenguajes,

entre ellos y solo por mencionar algunos, se encuentra Inglés, Chino, Español, Alemán y Árabe. Además, en su red neuronal cuenta con 12 capas, 768 unidades ocultas y 110 millones de parámetros.

Fuente: [Extending-Google-BERT-as-Question-and-Answering-model-and-Chatbot](#)