StruBERT: Structure-aware BERT for Table Search and Matching

Seminar “Modern Infomation Retrieval”, Summer 2023

Nicolas Hellthaler

Heidelberg University
Institute of Computer Science
nicolas.hellthaler@stud.uni-heidelberg.de

June 27, 2023

Outline

1. Motivation
2. Introduction
3. StruBERT Architecture
4. Evaluation
5. Reflection

About the Paper

StruBERT: Structure-aware BERT for Table Search and Matching [4]:

- Trabelsi, Chen, Zhang, Davison, Heflin
- presented at the 2022 WWW (now ACM Web Conference)

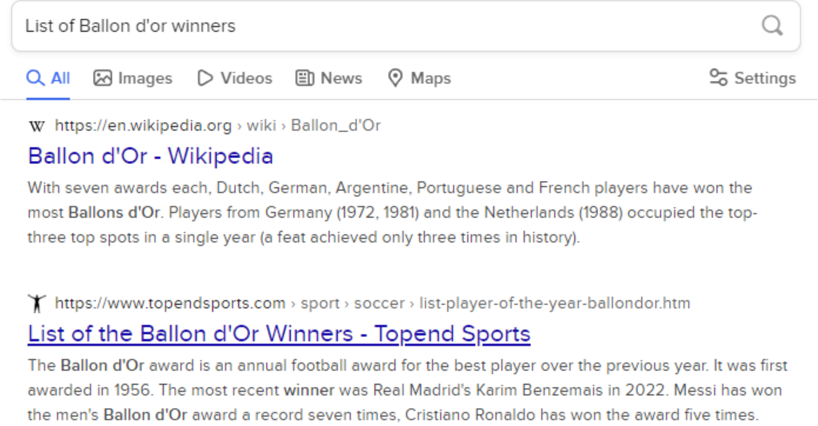
Contribution: New state-of-the-art model for…

- Table Search
- Table Matching

Outline

1. Motivation
2. Introduction
3. StruBERT Architecture
4. Evaluation
5. Reflection

Table Search



List of Ballon d'or winners

https://en.wikipedia.org › wiki › Ballon_d'Or

Ballon d'Or - Wikipedia

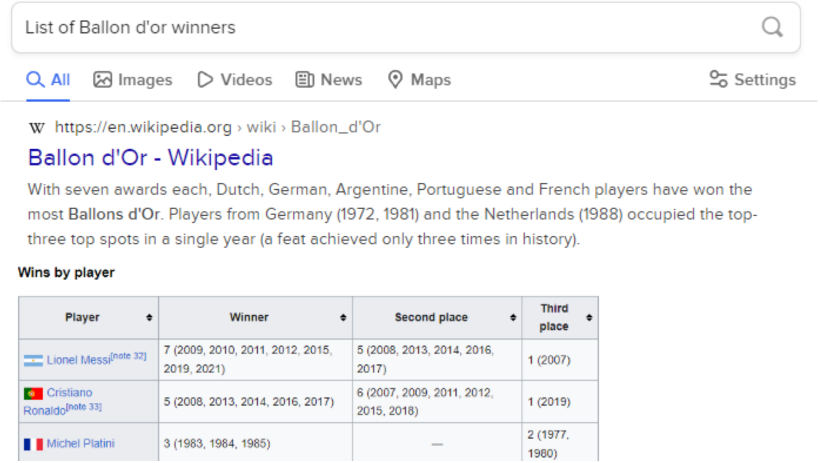
With seven awards each, Dutch, German, Argentine, Portuguese and French players have won the most **Ballons d'Or**. Players from Germany (1972, 1981) and the Netherlands (1988) occupied the top-three top spots in a single year (a feat achieved only three times in history).

https://www.topendsports.com › sport › soccer › list-player-of-the-year-ballondor.htm

List of the Ballon d'Or Winners - Topend Sports

The **Ballon d'Or** award is an annual football award for the best player over the previous year. It was first awarded in 1956. The most recent **winner** was Real Madrid's Karim Benzemais in 2022. Messi has won the men's **Ballon d'Or** award a record seven times, Cristiano Ronaldo has won the award five times.

Table Search



W https://en.wikipedia.org › wiki › Ballon_d'Or

Ballon d'Or - Wikipedia

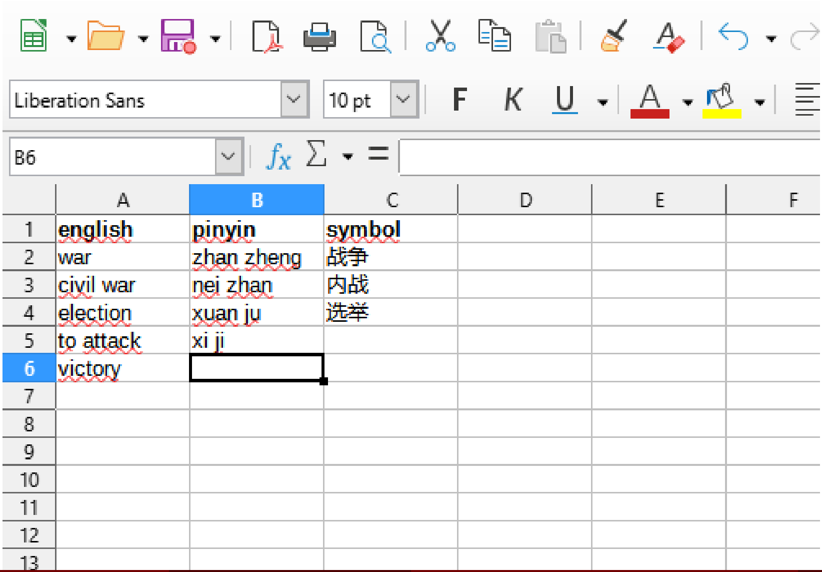
With seven awards each, Dutch, German, Argentine, Portuguese and French players have won the most **Ballons d'Or**. Players from Germany (1972, 1981) and the Netherlands (1988) occupied the top-three top spots in a single year (a feat achieved only three times in history).

Wins by player

Player	Winner	Second place	Third place
Lionel Messi[note 32]	7 (2009, 2010, 2011, 2012, 2015, 2019, 2021)	5 (2008, 2013, 2014, 2016, 2017)	1 (2007)
Cristiano Ronaldo[note 33]	5 (2008, 2013, 2014, 2016, 2017)	6 (2007, 2009, 2011, 2012, 2015, 2018)	1 (2019)
Michel Platini	3 (1983, 1984, 1985)	—	2 (1977, 1980)

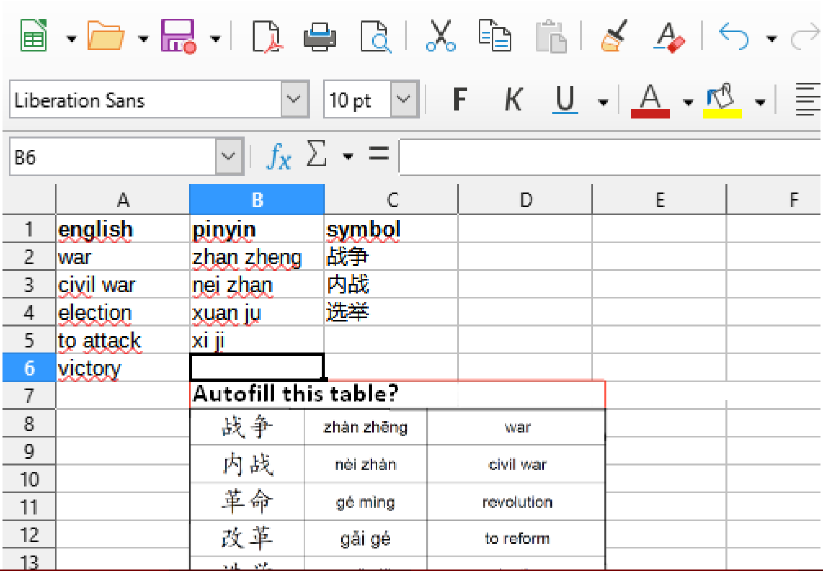
Figure: Ballon d'Or in Wikipedia, *Source: https://en.wikipedia.org/wiki/Ballon_d%27Or*

Table Matching



	A	B	C	D	E	F
1	english	pinyin	symbol			
2	war	zhan zheng	战争			
3	civil war	nei zhan	内战			
4	election	xuan ju	选举			
5	to attack	xi ji				
6	victory					
7						
8						
9						
10						
11						
12						
13						

Table Matching



	A	B	C	D	E	F
1	english	pinyin	symbol			
2	war	zhan zheng	战争			
3	civil war	nei zhan	内战			
4	election	xuan ju	选举			
5	to attack	xi ji				
6	victory					

Autofill this table?

战争	zhàn zhēng	war
内战	nèi zhàn	civil war
革命	gé mìng	revolution
改革	gǎi gé	to reform

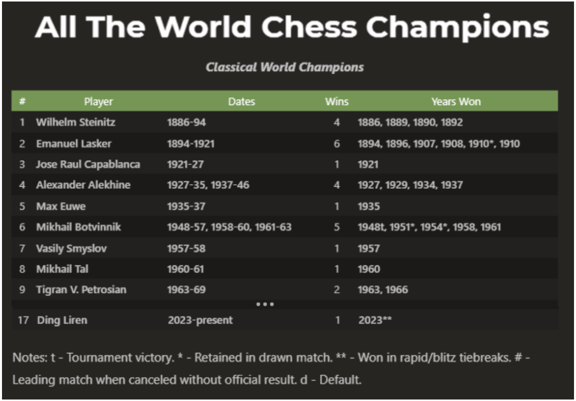
Outline

1. Motivation
2. Introduction
3. StruBERT Architecture
4. Evaluation
5. Reflection

A simple Table

$\mathbf{c_1}$	$\mathbf{c_2}$	$\cdots$	$\mathbf{c_l}$
v_{11}	v_{12}	$\cdots$	v_{1l}
v_{21}	v_{22}	$\cdots$	v_{2l}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$v_{(s-1)1}$	$v_{(s-1)2}$	$\cdots$	$v_{(s-1)l}$

Tables in practice



All The World Chess Champions

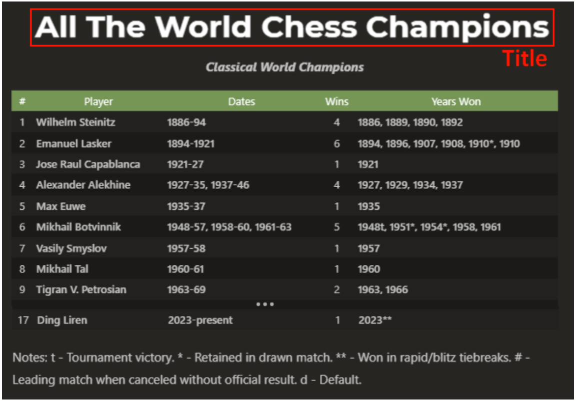
Classical World Champions

#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		•••		
17	Ding Liren	2023-present	1	2023**

Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

Figure: World Chess Champions on Chess.com, *Source: https://www.chess.com/article/view/world-chess-champions*

Tables in practice



All The World Chess Champions

Title

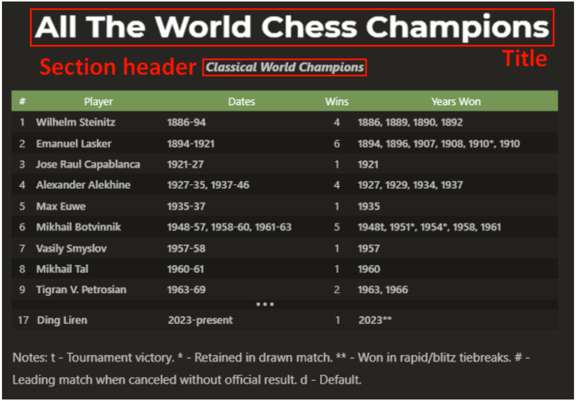
Classical World Champions

#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		...		
17	Ding Liren	2023-present	1	2023**

Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

Figure: World Chess Champions on Chess.com, *Source: https://www.chess.com/article/view/world-chess-champions*

Tables in practice



All The World Chess Champions

Title

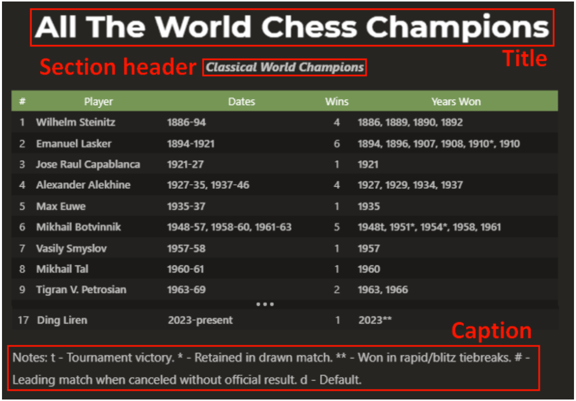
Section header Classical World Champions

#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		•••		
17	Ding Liren	2023-present	1	2023**

Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

Figure: World Chess Champions on Chess.com, *Source: https://www.chess.com/article/view/world-chess-champions*

Tables in practice



All The World Chess Champions

Title

Section header Classical World Champions

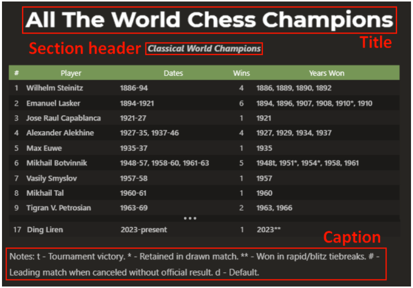
#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		...		
17	Ding Liren	2023-present	1	2023**

Caption

Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

Figure: World Chess Champions on Chess.com, *Source: https://www.chess.com/article/view/world-chess-champions*

Table Attributes



All The World Chess Champions

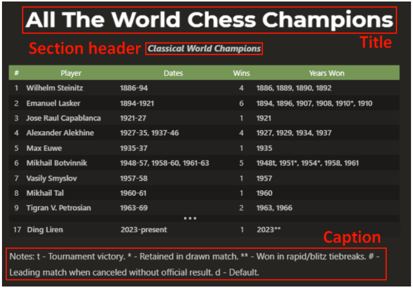
Classical World Champions

#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		...		
17	Ding Liren	2023-present	1	2023**

Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

- l Column headers: $c_1, c_2, \ldots, c_l$
- l Data types: $t_1, t_2, \ldots, t_l \in [real, text]$
- $(s-1)$ Data values per column: $v_{1i}, v_{2i}, \ldots, v_{(s-1)i}$
- p Related text fields: $f_1, f_2, \ldots, f_p$

Table Attributes


Title
Section header
Caption

All The World Chess Champions

Classical World Champions

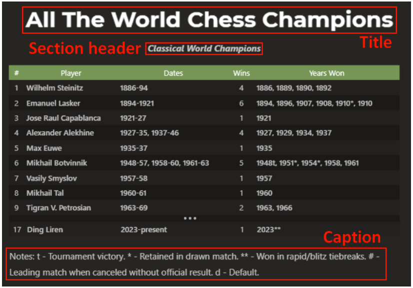
#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		...		
17	Ding Liren	2023-present	1	2023**

Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

- l Column headers: $c_1, c_2, \ldots, c_l$
- l Data types: $t_1, t_2, \ldots, t_l \in [real, text]$
- $(s-1)$ Data values per column: $v_{1i}, v_{2i}, \ldots, v_{(s-1)i}$
- p Related text fields: $f_1, f_2, \ldots, f_p$

$\Rightarrow$ Column headers + data values form structural information

Table Attributes



All The World Chess Champions

Classical World Champions

#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		•••		
17	Ding Liren	2023-present	1	2023**

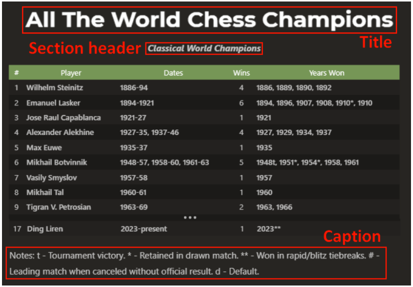
Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

- l Column headers: $c_1, c_2, \dots, c_l$
- l Data types: $t_1, t_2, \dots, t_l \in [real, text]$
- $(s-1)$ Data values per column: $v_{1i}, v_{2i}, \dots, v_{(s-1)i}$
- p Related text fields: $f_1, f_2, \dots, f_p$

⇒ Column headers + data values form structural information

⇒ Text fields form textual information

Processing Tables



All The World Chess Champions

Title

Section header Classical World Champions

#	Player	Dates	Wins	Years Won
1	Wilhelm Steinitz	1886-94	4	1886, 1889, 1890, 1892
2	Emanuel Lasker	1894-1921	6	1894, 1896, 1907, 1908, 1910*, 1910
3	Jose Raul Capablanca	1921-27	1	1921
4	Alexander Alekhine	1927-35, 1937-46	4	1927, 1929, 1934, 1937
5	Max Euwe	1935-37	1	1935
6	Mikhail Botvinnik	1948-57, 1958-60, 1961-63	5	1948t, 1951*, 1954*, 1958, 1961
7	Vasily Smyslov	1957-58	1	1957
8	Mikhail Tal	1960-61	1	1960
9	Tigran V. Petrosian	1963-69	2	1963, 1966
		•••		
17	Ding Liren	2023-present	1	2023**

Caption

Notes: t - Tournament victory. * - Retained in drawn match. ** - Won in rapid/blitz tiebreaks. # - Leading match when canceled without official result. d - Default.

What to use?

Ad Hoc Table Retrieval [6]:

- One core column
- Textual information

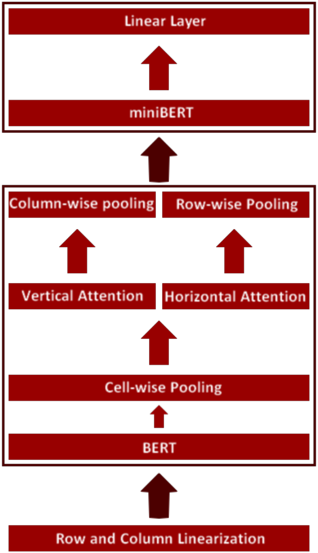
TabSim [3] / TaBERT [5]:

- All data cells
- BERT [1] to process text

Outline

1. Motivation
2. Introduction
3. StruBERT Architecture
4. Evaluation
5. Reflection

StruBERT


Linear Layer
miniBERT
Column-wise pooling
Row-wise Pooling
Vertical Attention
Horizontal Attention
Cell-wise Pooling
BERT
Row and Column Linearization

Converting Tables

Famous soccer players

Player	Team	Number
Ronaldo	Manchester United	7
Messi	Paris	30
Ramos	Real Madrid	4

Table: This table shows infomation about soccer players.

- $l = 3$ Column headers: $c_1 = \textit{Player}, c_2 = \textit{Team}, c_3 = \textit{Number}$
- $l = 3$ Data types: $t_1 = \textit{text}, t_2 = \textit{text}, t_3 = \textit{real}$
- $(s - 1) * l = 9$ Data values: v_{ij}
- $p = 2$ Related text fields:
  $f_1 = \text{"Famous soccer players"}, f_2 = \text{"This Table shows \ldots"}$

Column and Row Linearization

Player	Team	Number
Ronaldo	Manchester United	7
Messi	Paris	30
Ramos	Real Madrid	4

$$\tilde{c}_i = c_i t_i v_{1i}[\text{SEP}] c_i t_i v_{2i}[\text{SEP}] \ldots [\text{SEP}] c_i t_i v_{(s-1)i}[\text{SEP}]$$

$$\tilde{c}_1 = \text{Player text Ronaldo [SEP] Player text Messi [SEP]} \ldots$$

Column and Row Linearization

Player	Team	Number
Ronaldo	Manchester United	7
Messi	Paris	30
Ramos	Real Madrid	4

$$\tilde{c}_i = c_i t_i v_{1i}[\mathrm{SEP}] c_i t_i v_{2i}[\mathrm{SEP}] \ldots [\mathrm{SEP}] c_i t_i v_{(s-1)i}[\mathrm{SEP}]$$

$$\tilde{c}_1 = \text{Player text Ronaldo [SEP] Player text Messi [SEP]} \ldots$$

$$\tilde{r}_i = c_1 t_1 v_{i1}[\mathrm{SEP}] c_2 t_2 v_{i2}[\mathrm{SEP}] \ldots [\mathrm{SEP}] c_l t_l v_{il}[\mathrm{SEP}]$$

$$\tilde{r}_2 = \text{Player text Messi [SEP] Team text Paris [SEP]} \ldots$$

Column and Row Linearization

Player	**Team**	**Number**
Ronaldo	Manchester United	7
Messi	Paris	30
Ramos	Real Madrid	4

$$
\begin{aligned}
\tilde{c}_i &= c_i t_i v_{1i}[\mathrm{SEP}] c_i t_i v_{2i}[\mathrm{SEP}] \ldots [\mathrm{SEP}] c_i t_i v_{(s-1)i}[\mathrm{SEP}] \\
\tilde{c}_1 &= \text{Player text Ronaldo [SEP] Player text Messi [SEP]} \ldots \\
\tilde{r}_i &= c_1 t_1 v_{i1}[\mathrm{SEP}] c_2 t_2 v_{i2}[\mathrm{SEP}] \ldots [\mathrm{SEP}] c_l t_l v_{il}[\mathrm{SEP}] \\
\tilde{r}_2 &= \text{Player text Messi [SEP] Team text Paris [SEP]} \ldots
\end{aligned}
$$

Textual information missing!

Adding the Textual Information

How do we integrate f_1 (page title) and f_2 (caption)?

Adding the Textual Information

How do we integrate f_1 (page title) and f_2 (caption)?
$\Rightarrow$ Simply use as prefix

$$\bar{c}_i = \text{[CLS]} f_1 \text{[SEP]} f_2 \text{[SEP]} \ldots \text{[SEP]} f_p \text{[SEP]} \tilde{c}_i \text{[SEP]}$$
$$\bar{c}_1 = \text{[CLS]Famous Soccer Players[SEP]This Table shows } \ldots \text{[SEP]} \tilde{c}_1 \text{[SEP]}$$

Adding the Textual Information

How do we integrate f_1 (page title) and f_2 (caption)?
⇒ Simply use as prefix

$$\bar{c}_i = \text{[CLS]} f_1 \text{[SEP]} f_2 \text{[SEP]} \ldots \text{[SEP]} f_p \text{[SEP]} \tilde{c}_i \text{[SEP]}$$
$$\bar{c}_1 = \text{[CLS]Famous Soccer Players[SEP]This Table shows } \ldots \text{[SEP]} \tilde{c}_1 \text{[SEP]}$$

$$\bar{r}_i = \text{[CLS]} f_1 \text{[SEP]} f_2 \text{[SEP]} \ldots \text{[SEP]} f_p \text{[SEP]} \tilde{r}_i \text{[SEP]}$$
$$\bar{r}_2 = \text{[CLS]Famous Soccer Players[SEP]This Table shows } \ldots \text{[SEP]} \tilde{r}_2 \text{[SEP]}$$

Adding the Textual Information

How do we integrate f_1 (page title) and f_2 (caption)?
$\Rightarrow$ Simply use as prefix

$\bar{c}_i = \text{[CLS]} f_1 \text{[SEP]} f_2 \text{[SEP]} \dots \text{[SEP]} f_p \text{[SEP]} \tilde{c}_i \text{[SEP]}$
$\bar{c}_1 = \text{[CLS]Famous Soccer Players[SEP]This Table shows} \dots \text{[SEP]} \tilde{c}_1 \text{[SEP]}$

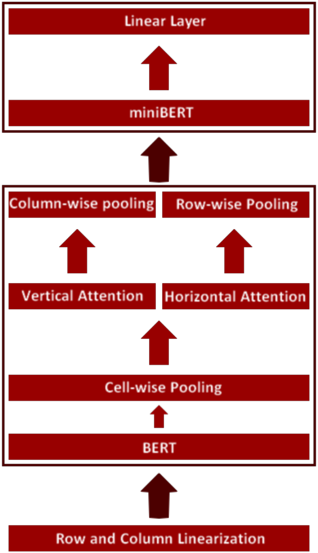
$\bar{r}_i = \text{[CLS]} f_1 \text{[SEP]} f_2 \text{[SEP]} \dots \text{[SEP]} f_p \text{[SEP]} \tilde{r}_i \text{[SEP]}$
$\bar{r}_2 = \text{[CLS]Famous Soccer Players[SEP]This Table shows} \dots \text{[SEP]} \tilde{r}_2 \text{[SEP]}$

This is new!

$$\bar{\mathcal{C}} = \{\bar{c}_1, \bar{c}_2, \dots, \bar{c}_l\}$$

$$\bar{\mathcal{R}} = \{\bar{r}_1, \bar{r}_2, \dots, \bar{r_{(s-1)}}\}$$

StruBERT


Linear Layer
miniBERT
Column-wise pooling
Row-wise Pooling
Vertical Attention
Horizontal Attention
Cell-wise Pooling
BERT
Row and Column Linearization

BERT

$$\text{[CLS]}\,\tilde{T}_{ej}\text{[SEP]}c_i t_i v_{1i}\text{[SEP]}c_i t_i v_{2i}\text{[SEP]}\dots\text{[SEP]}c_i t_i v_{(s-1)i}\text{[SEP]}$$

$$\text{[CLS]}\,\tilde{T}_{ej}\text{[SEP]Player text Ronaldo [SEP] Player text Messi [SEP]}\dots$$

BERT

BERT

$$[\text{CLS}]\,\tilde{T}_{ej}[\text{SEP}]c_i t_i v_{1i}[\text{SEP}]c_i t_i v_{2i}[\text{SEP}]\ldots[\text{SEP}]c_i t_i v_{(s-1)i}[\text{SEP}]$$

$$[\text{CLS}]\,\tilde{T}_{ej}[\text{SEP}]\text{Player text Ronaldo [SEP] Player text Messi [SEP]}\ldots$$

BERT

$$\boldsymbol{[\text{CLS}]\,\tilde{T}_{ej}[\text{SEP}]c_i t_i v_{1i}[\text{SEP}]c_i t_i v_{2i}[\text{SEP}]\ldots[\text{SEP}]c_i t_i v_{(s-1)i}[\text{SEP}]}$$

$$\boldsymbol{[\text{CLS}]\,\tilde{T}_{ej}[\text{SEP}]\text{Player text Ronaldo [SEP] Player text Messi [SEP]}\ldots}$$

Transformer Count:

Average pooling

$$\text{[CLS]}\,\tilde{\boldsymbol{T}}_{\boldsymbol{ej}}\text{[SEP]}\,\underbrace{\text{Player text Ronaldo}}_{c_1 t_1 v_{11}}\,\text{[SEP]}\,\underbrace{\text{Player text Messi}}_{c_1 t_1 v_{21}}\text{[SEP]}\ldots$$

$$\boldsymbol{v_{ki}} = \frac{\sum\limits_{w \in BertTok(c_i t_i v_{ki})} BERT(w)}{|BertTok(c_i t_i v_{ki})|}$$

$$\boldsymbol{v_{11}} = \frac{\sum\limits_{w \in BertTok(\text{Player text Ronaldo})} BERT(w)}{|BertTok(\text{Player text Ronaldo})|}$$

Average pooling

$$[\text{CLS}]\,\tilde{T}_{ej}[\text{SEP}]\,\underbrace{\text{Player text Ronaldo}}_{c_1 t_1 v_{11}}\,[\text{SEP}]\,\underbrace{\text{Player text Messi}}_{c_1 t_1 v_{21}}[\text{SEP}]\,\ldots$$

$$\boldsymbol{v_{ki}} = \frac{\sum\limits_{w \in BertTok(c_i t_i v_{ki})} BERT(w)}{|BertTok(c_i t_i v_{ki})|}$$

$$\boldsymbol{v_{11}} = \frac{\sum\limits_{w \in BertTok(\text{Player text Ronaldo})} BERT(w)}{|BertTok(\text{Player text Ronaldo})|}$$

$$= \frac{BERT(Player) + BERT(text) + BERT(Ronaldo)}{3}$$

Attention Please!

Vertical Self-Attention + Column-wise Pooling

$\bar{r}_1$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{11}	[SEP]	v_{12}	[SEP]	v_{13}
$\bar{r}_2$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{21}	[SEP]	v_{22}	[SEP]	v_{23}
$\bar{r}_3$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{31}	[SEP]	v_{32}	[SEP]	v_{33}

Attention Please!

Vertical Self-Attention + Column-wise Pooling

$\bar{r}_1$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{11}	[SEP]	v_{12}	[SEP]	v_{13}
$\bar{r}_2$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{21}	[SEP]	v_{22}	[SEP]	v_{23}
$\bar{r}_3$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{31}	[SEP]	v_{32}	[SEP]	v_{33}

$[\hat{CLS}]$	$\hat{v_{11}}$	$\hat{v_{12}}$	$\hat{v_{13}}$
$[\hat{CLS}]$	$\hat{v_{21}}$	$\hat{v_{22}}$	$\hat{v_{23}}$
$[\hat{CLS}]$	$\hat{v_{31}}$	$\hat{v_{32}}$	$\hat{v_{33}}$

$[CLS]_c$	c_1	c_2	c_3

Attention Please!

Vertical Self-Attention + Column-wise Pooling

$$\begin{array}{l|llllllll}
\bar{r}_1 & [\mathrm{CLS}] & \tilde{T}_{ej} & [\mathrm{SEP}] & v_{11} & [\mathrm{SEP}] & v_{12} & [\mathrm{SEP}] & v_{13} \\
\bar{r}_2 & [\mathrm{CLS}] & \tilde{T}_{ej} & [\mathrm{SEP}] & v_{21} & [\mathrm{SEP}] & v_{22} & [\mathrm{SEP}] & v_{23} \\
\bar{r}_3 & [\mathrm{CLS}] & \tilde{T}_{ej} & [\mathrm{SEP}] & v_{31} & [\mathrm{SEP}] & v_{32} & [\mathrm{SEP}] & v_{33}
\end{array}$$

$$\begin{array}{cccc}
[C\hat{L}S] & \hat{v_{11}} & \hat{v_{12}} & \hat{v_{13}} \\
[C\hat{L}S] & \hat{v_{21}} & \hat{v_{22}} & \hat{v_{23}} \\
[C\hat{L}S] & \hat{v_{31}} & \hat{v_{32}} & \hat{v_{33}}
\end{array}$$

$$\begin{array}{cccc}
[CLS]_c & c_1 & c_2 & c_3
\end{array}$$

- 1 Column guided [CLS] embedding
- l Column embeddings

Attention Please!

Vertical Self-Attention + Column-wise Pooling

$\bar{r}_1$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{11}	[SEP]	v_{12}	[SEP]	v_{13}
$\bar{r}_2$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{21}	[SEP]	v_{22}	[SEP]	v_{23}
$\bar{r}_3$	[CLS]	$\tilde{T}_{ej}$	[SEP]	v_{31}	[SEP]	v_{32}	[SEP]	v_{33}

$[C\hat{L}S]$	$\hat{v}_{11}$	$\hat{v}_{12}$	$\hat{v}_{13}$
$[C\hat{L}S]$	$\hat{v}_{21}$	$\hat{v}_{22}$	$\hat{v}_{23}$
$[C\hat{L}S]$	$\hat{v}_{31}$	$\hat{v}_{32}$	$\hat{v}_{33}$

$[\mathrm{CLS}]_c$	c_1	c_2	c_3

- 1 Column guided [CLS] embedding
- l Column embeddings

Similar to TaBERT

Attention Please! / 2

Horizontal Self-Attention + Row-wise Pooling

$\bar{c}_1$	$\bar{c}_2$	$\bar{c}_3$
[CLS]	[CLS]	[CLS]
$\tilde{T}_{ej}$	$\tilde{T}_{ej}$	$\tilde{T}_{ej}$
[SEP]	[SEP]	[SEP]
v_{11}	v_{12}	v_{13}
[SEP]	[SEP]	[SEP]
v_{21}	v_{22}	v_{23}
[SEP]	[SEP]	[SEP]
v_{31}	v_{32}	v_{33}

Attention Please! / 2

Horizontal Self-Attention + Row-wise Pooling

$\bar{c}_1$	$\bar{c}_2$	$\bar{c}_3$
[CLS]	[CLS]	[CLS]
$\tilde{T}_{ej}$	$\tilde{T}_{ej}$	$\tilde{T}_{ej}$
[SEP]	[SEP]	[SEP]
v_{11}	v_{12}	v_{13}
[SEP]	[SEP]	[SEP]
v_{21}	v_{22}	v_{23}
[SEP]	[SEP]	[SEP]
v_{31}	v_{32}	v_{33}

$[\hat{CLS}]$	$[\hat{CLS}]$	$[\hat{CLS}]$
$\hat{v}_{11}$	$\hat{v}_{12}$	$\hat{v}_{13}$
$\hat{v}_{21}$	$\hat{v}_{22}$	$\hat{v}_{23}$
$\hat{v}_{31}$	$\hat{v}_{32}$	$\hat{v}_{33}$

Attention Please! / 2

Horizontal Self-Attention + Row-wise Pooling

$\bar{c}_1$	$\bar{c}_2$	$\bar{c}_3$
[CLS]	[CLS]	[CLS]
$\tilde{T}_{ej}$	$\tilde{T}_{ej}$	$\tilde{T}_{ej}$
[SEP]	[SEP]	[SEP]
v_{11}	v_{12}	v_{13}
[SEP]	[SEP]	[SEP]
v_{21}	v_{22}	v_{23}
[SEP]	[SEP]	[SEP]
v_{31}	v_{32}	v_{33}

$[C\hat{L}S]$	$[C\hat{L}S]$	$[C\hat{L}S]$
$\hat{v}_{11}$	$\hat{v}_{12}$	$\hat{v}_{13}$
$\hat{v}_{21}$	$\hat{v}_{22}$	$\hat{v}_{23}$
$\hat{v}_{31}$	$\hat{v}_{32}$	$\hat{v}_{33}$

$[\text{CLS}]_r$
r_1
r_2
r_3

Attention Please! / 2

Horizontal Self-Attention + Row-wise Pooling

$\bar{c}_1$	$\bar{c}_2$	$\bar{c}_3$
[CLS]	[CLS]	[CLS]
$\tilde{T}_{ej}$	$\tilde{T}_{ej}$	$\tilde{T}_{ej}$
[SEP]	[SEP]	[SEP]
v_{11}	v_{12}	v_{13}
[SEP]	[SEP]	[SEP]
v_{21}	v_{22}	v_{23}
[SEP]	[SEP]	[SEP]
v_{31}	v_{32}	v_{33}

$[\hat{CLS}]$	$[\hat{CLS}]$	$[\hat{CLS}]$
$\hat{v_{11}}$	$\hat{v_{12}}$	$\hat{v_{13}}$
$\hat{v_{21}}$	$\hat{v_{22}}$	$\hat{v_{23}}$
$\hat{v_{31}}$	$\hat{v_{32}}$	$\hat{v_{33}}$

$[\text{CLS}]_r$
r_1
r_2
r_3

- 1 Row guided [CLS] embedding
- $s - 1$ Row embeddings

Transformer Count:

StruBERT Output

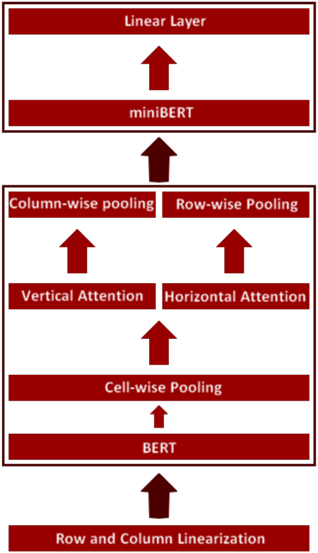
Player	Team	Number
Ronaldo	Manchester United	7
Messi	Paris	30
Ramos	Real Madrid	4

Table: T_i

$$StruBERT(T_i) = (\boldsymbol{E_r^i}, \boldsymbol{E_c^i}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{i}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{i}})$$

- $\boldsymbol{E_r^i}$: $s-1$ Row embeddings
- $\boldsymbol{E_c^i}$: l Column embeddings
- $\textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{i}}$: 1 Row guided [CLS] embedding
- $\textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{i}}$: 1 Column guided [CLS] embedding

StruBERT

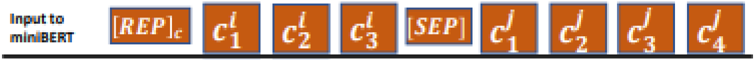

Linear Layer
miniBERT
Column-wise pooling
Row-wise Pooling
Vertical Attention
Horizontal Attention
Cell-wise Pooling
BERT
Row and Column Linearization

StruBERT in Action: Table Matching

1. Apply StruBERT to both Tables:
$StruBERT(T_i) = (\boldsymbol{E_r^i}, \boldsymbol{E_c^i}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{i}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{i}})$
$StruBERT(T_j) = (\boldsymbol{E_r^j}, \boldsymbol{E_c^j}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{j}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{j}})$

StruBERT in Action: Table Matching

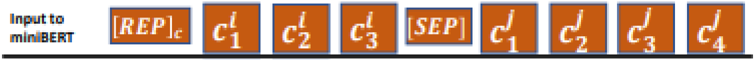
1. Apply StruBERT to both Tables:
   $StruBERT(T_i) = (\boldsymbol{E_r^i}, \boldsymbol{E_c^i}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{i}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{i}})$
   $StruBERT(T_j) = (\boldsymbol{E_r^j}, \boldsymbol{E_c^j}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{j}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{j}})$
2. Input row and column embeddings to miniBERT:



$\Rightarrow$ miniBERT is a new ranking model!

StruBERT in Action: Table Matching

1. Apply StruBERT to both Tables:
   $StruBERT(T_i) = (\boldsymbol{E_r^i}, \boldsymbol{E_c^i}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{i}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{i}})$
   $StruBERT(T_j) = (\boldsymbol{E_r^j}, \boldsymbol{E_c^j}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{j}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{j}})$
2. Input row and column embeddings to miniBERT:



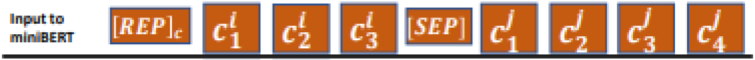
$\Rightarrow$ miniBERT is a new ranking model!

3. Build final output:

$$\textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{i}} \odot \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{j}} \oplus \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{i}} \odot \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{j}} \oplus miniBERT(\textbf{[REP]}_{\boldsymbol{r}}) \oplus miniBERT(\textbf{[REP]}_{\boldsymbol{c}})$$

StruBERT in Action: Table Matching

1. Apply StruBERT to both Tables:
   $$StruBERT(T_i) = (\boldsymbol{E_r^i}, \boldsymbol{E_c^i}, \mathbf{[CLS]_r^i}, \mathbf{[CLS]_c^i})$$
   $$StruBERT(T_j) = (\boldsymbol{E_r^j}, \boldsymbol{E_c^j}, \mathbf{[CLS]_r^j}, \mathbf{[CLS]_c^j})$$
2. Input row and column embeddings to miniBERT:



Input to miniBERT $[REP]_c$ c_1^i c_2^i c_3^i $[SEP]$ c_1^j c_2^j c_3^j c_4^j

⇒ miniBERT is a new ranking model!

3. Build final output:

$$\mathbf{[CLS]_r^i} \odot \mathbf{[CLS]_r^j} \oplus \mathbf{[CLS]_c^i} \odot \mathbf{[CLS]_c^j} \oplus miniBERT(\mathbf{[REP]_r}) \oplus miniBERT(\mathbf{[REP]_c})$$

Transformer Count:

StruBERT in Action: Table Search

1. Insert query-keywords $q_1, q_2, \ldots, q_m$ into row and column sequences:
   - $\tilde{c}_1 = \text{Player text Ronaldo [SEP] Player text Messi [SEP]} \ldots$

StruBERT in Action: Table Search

1. Insert query-keywords $q_1, q_2, \ldots, q_m$ into row and column sequences:
    - $\tilde{c}_1 = \text{Player text Ronaldo [SEP] Player text Messi [SEP]} \ldots$
    - $\tilde{c}_1 = \textit{Title} \text{ [SEP] Player text Ronaldo [SEP] Player text Messi} \ldots$

StruBERT in Action: Table Search

1. Insert query-keywords $q_1, q_2, \ldots, q_m$ into row and column sequences:
   - $\tilde{c}_1 = \text{Player text Ronaldo [SEP] Player text Messi [SEP]} \ldots$
   - $\tilde{c}_1 = \textit{Title} \text{ [SEP] Player text Ronaldo [SEP] Player text Messi} \ldots$
   - $\tilde{c}_1 = \textit{Query} \text{ [SEP] } \textit{Title} \text{ [SEP] Player text Ronaldo [SEP] Player} \ldots$

StruBERT in Action: Table Search

1. Insert query-keywords $q_1, q_2, \ldots, q_m$ into row and column sequences:
   - $\tilde{c}_1 =$ Player text Ronaldo [SEP] Player text Messi [SEP] …
   - $\tilde{c}_1 =$ *Title* [SEP] Player text Ronaldo [SEP] Player text Messi …
   - $\tilde{c}_1 =$ *Query* [SEP] *Title* [SEP] Player text Ronaldo [SEP] Player …

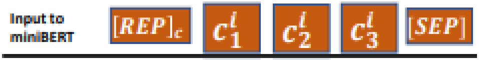
2. Apply StruBERT:
   $StruBERT(T_i) = (\boldsymbol{E_r^i(q)}, \boldsymbol{E_c^i(q)}, \textbf{[CLS]}_{\boldsymbol{r}}^{\boldsymbol{i}}, \textbf{[CLS]}_{\boldsymbol{c}}^{\boldsymbol{i}})$

StruBERT in Action: Table Search

1. Insert query-keywords $q_1, q_2, \ldots, q_m$ into row and column sequences:
   - $\tilde{c}_1$ = Player text Ronaldo [SEP] Player text Messi [SEP] ...
   - $\tilde{c}_1$ = *Title* [SEP] Player text Ronaldo [SEP] Player text Messi ...
   - $\tilde{c}_1$ = *Query* [SEP] *Title* [SEP] Player text Ronaldo [SEP] Player ...

2. Apply StruBERT:
   $StruBERT(T_i) = (\boldsymbol{E_r^i(q)}, \boldsymbol{E_c^i(q)}, [\mathbf{CLS}]_{\boldsymbol{r}}^{\boldsymbol{i}}, [\mathbf{CLS}]_{\boldsymbol{c}}^{\boldsymbol{i}})$

3. Apply miniBERT:

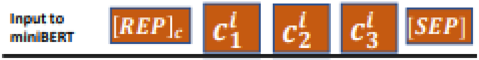


StruBERT in Action: Table Search

1. Insert query-keywords $q_1, q_2, \ldots, q_m$ into row and column sequences:
   - $\tilde{c}_1$ = Player text Ronaldo [SEP] Player text Messi [SEP] . . .
   - $\tilde{c}_1$ = *Title* [SEP] Player text Ronaldo [SEP] Player text Messi . . .
   - $\tilde{c}_1$ = *Query* [SEP] *Title* [SEP] Player text Ronaldo [SEP] Player . . .

2. Apply StruBERT:
   $StruBERT(T_i) = (\boldsymbol{E_r^i(q)}, \boldsymbol{E_c^i(q)}, \mathbf{[CLS]_r^i}, \mathbf{[CLS]_c^i})$

3. Apply miniBERT:



4. Build final output:

$$\mathbf{[CLS]_r^i} \oplus \mathbf{[CLS]_c^i} \oplus miniBERT(\mathbf{[REP]_r}) \oplus miniBERT(\mathbf{[REP]_c})$$

Outline

1. Motivation
2. Introduction
3. StruBERT Architecture
4. Evaluation
5. Reflection

Table Similarity: Datasets and Metrics

PMC:

- From scientific papers
- Tables + captions
- Tables as pairs with binary labels
- 1391 pairs

WikiTables:

- Wikipedia tables
- Tables + captions + page title + section title + column headings
- Tables as pairs with binary labels
- ca. 3000 pairs

Table Similarity: Datasets and Metrics

PMC:

- From scientific papers
- Tables + captions
- Tables as pairs with binary labels
- 1391 pairs

WikiTables:

- Wikipedia tables
- Tables + captions + page title + section title + column headings
- Tables as pairs with binary labels
- ca. 3000 pairs

5 fold cross-validation ⇒ macro-averaged metrics

Table Similarity: Results

Method Name	Macro-P	Macro-R	Macro-F	Accur.
Tfidf + MLP	0.7834	0.6735	0.6529	0.6951
TaBERT	0.9109	0.9024	0.9055	0.9067
StruBERT (CNN)	0.9293	0.9164	0.9205	0.9224
StruBERT	**0.9321**	**0.9284**	**0.9300**	**0.9310**

(a) PMC

Table Similarity: Results

Method Name	Macro-P	Macro-R	Macro-F	Accur.
Tfidf + MLP	0.7834	0.6735	0.6529	0.6951
TaBERT	0.9109	0.9024	0.9055	0.9067
StruBERT (CNN)	0.9293	0.9164	0.9205	0.9224
StruBERT	**0.9321**	**0.9284**	**0.9300**	**0.9310**

(a) PMC

Method Name	Macro-P	Macro-R	Macro-F	Accur.
Tfidf + MLP	0.6256	0.5022	0.3559	0.5378
TaBERT	0.9696	0.9626	0.9649	0.9653
StruBERT (CNN)	0.9782	0.9737	0.9753	0.9756
StruBERT	**0.9945**	**0.9938**	**0.9941**	**0.9942**

(b) WikiTables

Content-based Table Retrieval: Datasets and Metrics

Query by Example Data [7]:

- Adaptation of WikiTables
- 50 query-tables from different domains
- Tables as pairs with label:
    - 2 - highly relevant
    - 1 - relevant
    - 0 - irrelevant
- 2850 pairs

Content-based Table Retrieval: Datasets and Metrics

Query by Example Data [7]:

- Adaptation of WikiTables
- 50 query-tables from different domains
- Tables as pairs with label:
  - 2 - highly relevant
  - 1 - relevant
  - 0 - irrelevant
- 2850 pairs

information retrieval system ⇒ NDCG, MRR, MAP

Content-based Table Retrieval: Results

Method Name	NDCG@5	MRR	MAP
BM25	0.5369	0.5832	0.5417
TaBERT	0.5877	0.6120	0.5942
StruBERT (CNN)	0.6177	0.6378	0.6179
StruBERT	**0.6345**	**0.6601**	**0.6297**

Table: Query by Example Dataset

Keyword-based Table Retrieval: Datasets and Metrics

WikiTables:

- Wikipedia tables
- 60 natural language queries
- Table-query pairs with label:
  - 2 - highly relevant
  - 1 - relevant
  - 0 - irrelevant
- 3117 pairs

information retrieval system ⇒ NDCG, MRR, MAP

Evaluation: Keyword-based Table Retrieval

Method Name	NDCG@5	MRR	MAP
MultiField-BM25	0.4365	0.4882	0.4596
TaBERT	0.6055	0.6436	0.6146
StruBERT	**0.6393**	**0.6688**	**0.6378**

Table: WikiTables

Conclusion

Key Takeaways

- Early interactions between text and structure are important

Conclusion

Key Takeaways

- Early interactions between text and structure are important
- Attention = good

Conclusion

Key Takeaways

- Early interactions between text and structure are important
- Attention = good
- More attention = More good

Outline

1. Motivation
2. Introduction
3. StruBERT Architecture
4. Evaluation
5. Reflection

My Thoughts on the paper

I liked:

- Very understandably written
- Easy code access (and execution)

I did not like:

- Missing performance information

Sources I

J. Devlin, M. Chang, K. Lee, and K. Toutanova.
BERT: pre-training of deep bidirectional transformers for language understanding.
CoRR, abs/1810.04805, 2018.

A. Dhinakaran.
Demystifying ndcg.
https://towardsdatascience.com/demystifying-ndcg-bee3be58cfe0, 2023.

M. Habibi, J. Starlinger, and U. Leser.
Tabsim: A siamese neural network for accurate estimation of table similarity.
In *2020 IEEE International Conference on Big Data (Big Data)*, pages 930–937, 2020.

M. Trabelsi, Z. Chen, S. Zhang, B. D. Davison, and J. Heflin.
Strubert: Structure-aware bert for table search and matching.
In *Proceedings of the ACM Web Conference 2022*, WWW '22, page 442–451, New York, NY, USA, 2022. Association for Computing Machinery.

P. Yin, G. Neubig, W. Yih, and S. Riedel.
Tabert: Pretraining for joint understanding of textual and tabular data.
CoRR, abs/2005.08314, 2020.

S. Zhang and K. Balog.
Ad hoc table retrieval using semantic similarity.
In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, page 1553–1562, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee.

Sources II

S. Zhang and K. Balog.
Recommending related tables.
CoRR, abs/1907.03595, 2019.

Questions

Questions

Example of PMC data

Gene	Forward	Reverse
Gapdh	ACCAAATCCGTTGACTCCGAC	TTCGACAGTCAGCCGCATCT
Gpr40	AGTGTGGTGCTTAATCCGCT	AGTGGCGTTACTTCTGGGAC
E-cadherin	CTTGGAGCCGCAGCCTCT	ACACCATCTGTGCCCACTTT
Beta-catenin	ACGGAGGAAGGTCTGAGGAG	GCCGCTTTTCTGTCTGGTTC

Table: Primer sequences for in vitro experiments. [3]

Mentioned Metrics I

Normalized Discounted Cumulative Gain [2]

$$NDCG@K = \frac{DCG@K}{IDCG@K} = \frac{\sum_{i=1}^{k\ (actual\ order)} \frac{Gains}{log_2(i+1)}}{\sum_{i=1}^{k\ (ideal\ order)} \frac{Gains}{log_2(i+1)}}$$

Mean Average Precison

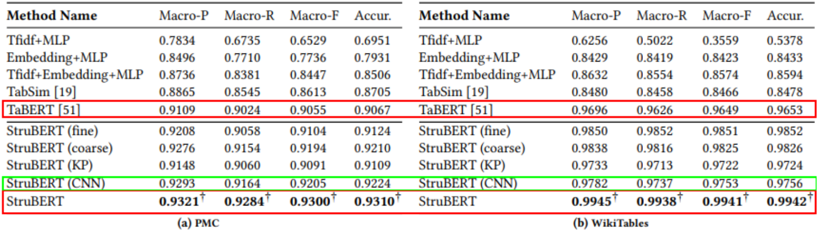
$$mAP = \frac{1}{|Q|} \sum_{q=1}^{|Q|} AveP(q)$$

Mentioned Metrics II

Mean reciprocal rank

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i}$$

Table Similarity Evaluation



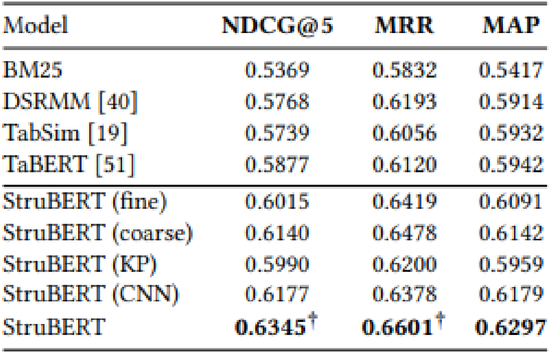
Method Name	Macro-P	Macro-R	Macro-F	Accur.
Tfidf+MLP	0.7834	0.6735	0.6529	0.6951
Embedding+MLP	0.8496	0.7710	0.7736	0.7931
Tfidf+Embedding+MLP	0.8736	0.8381	0.8447	0.8506
TabSim [19]	0.8865	0.8545	0.8613	0.8705
TaBERT [51]	0.9109	0.9024	0.9055	0.9067
StruBERT (fine)	0.9208	0.9058	0.9104	0.9124
StruBERT (coarse)	0.9276	0.9154	0.9194	0.9210
StruBERT (KP)	0.9148	0.9060	0.9091	0.9109
StruBERT (CNN)	0.9293	0.9164	0.9205	0.9224
StruBERT	**0.9321**†	**0.9284**†	**0.9300**†	**0.9310**†

(a) PMC

Method Name	Macro-P	Macro-R	Macro-F	Accur.
Tfidf+MLP	0.6256	0.5022	0.3559	0.5378
Embedding+MLP	0.8429	0.8419	0.8423	0.8433
Tfidf+Embedding+MLP	0.8632	0.8554	0.8574	0.8594
TabSim [19]	0.8480	0.8458	0.8466	0.8478
TaBERT [51]	0.9696	0.9626	0.9649	0.9653
StruBERT (fine)	0.9850	0.9852	0.9851	0.9852
StruBERT (coarse)	0.9838	0.9816	0.9825	0.9826
StruBERT (KP)	0.9733	0.9713	0.9722	0.9724
StruBERT (CNN)	0.9782	0.9737	0.9753	0.9756
StruBERT	**0.9945**†	**0.9938**†	**0.9941**†	**0.9942**†

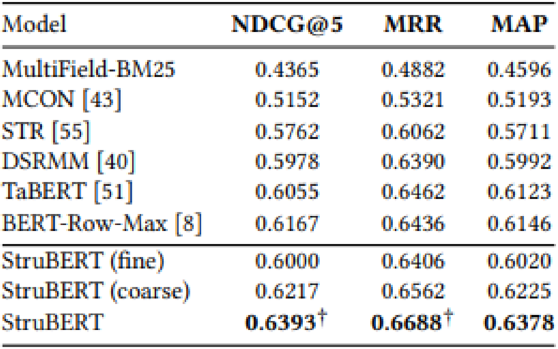
(b) WikiTables

Content-based Table Retrieval Evaluation



Model	**NDCG@5**	**MRR**	**MAP**
BM25	0.5369	0.5832	0.5417
DSRMM [40]	0.5768	0.6193	0.5914
TabSim [19]	0.5739	0.6056	0.5932
TaBERT [51]	0.5877	0.6120	0.5942
StruBERT (fine)	0.6015	0.6419	0.6091
StruBERT (coarse)	0.6140	0.6478	0.6142
StruBERT (KP)	0.5990	0.6200	0.5959
StruBERT (CNN)	0.6177	0.6378	0.6179
StruBERT	**0.6345**†	**0.6601**†	**0.6297**

Keyword-based Table Retrieval Evaluation



Model	**NDCG@5**	**MRR**	**MAP**
MultiField-BM25	0.4365	0.4882	0.4596
MCON [43]	0.5152	0.5321	0.5193
STR [55]	0.5762	0.6062	0.5711
DSRMM [40]	0.5978	0.6390	0.5992
TaBERT [51]	0.6055	0.6462	0.6123
BERT-Row-Max [8]	0.6167	0.6436	0.6146
StruBERT (fine)	0.6000	0.6406	0.6020
StruBERT (coarse)	0.6217	0.6562	0.6225
StruBERT	**0.6393**†	**0.6688**†	**0.6378**