

Action prediction . . .

Giovanni Cravero
Politecnico di Torino

s193598@studenti.polito.it

Eugenio Dosualdo
Politecnico di Torino

s271880@studenti.polito.it

Marco Rondina
Politecnico di Torino

s280096@studenti.polito.it

Abstract

Last but not least (this is the last piece of work)

1. Introduction

Brief description of the task and the goals.

2. Background

Introduction

2.1. SIMMC

2.1.1 Data description

2.2. BERT and Transformers

3. Model

What and why

3.1. Input manipulation

Here we can describe the tensor dataset structure and the tokenized input value, *e.g.*

[CLS]	Q1	L1			
[CLS]	Q1	A1	[SEP]	Q2	L2
[CLS]	Q2	A2	[SEP]	Q3	L3
[CLS]	Q3	A3	[SEP]	Q4	L4

Table 1. Sentences composition

The tab. 2 . . .

3.2. Added layers - Activation functions

3.3. Loss function

3.3.1 Actions

3.3.2 Attributes

3.4. Tuning

- epochs
- batches

- learning rate

- . . .

4. Results

4.1. Evaluation on test set

Inserire train loss ecc.?

Batch	Epoch	Out Hid.	Act. acc.	Perpl.	Att. acc.	Tr.Loss
14	2	256	84.55	2.09	71.05	
14	2	768	84.97	2.24	73.16	
14	3	256	84.91	2.40	72.50	
14	3	768	84.68	2.43	72.87	
12	4	256	84.25	3.16	73.70	
14	4	256	85.21	2.57	74.70	

Table 2. Result table

DA PROVARE batch 14 con 4 epoche e 768 livelli di output nell'intermedio

overfitting analysis - plots - confusion matrix

5. Conclusion