

Deep Comedy

Lorenzo Mario Amorosa - 0000948133

Sami Mohammed Osman - 0000929408

Index

- Data Processing
- Subword-Level Tokenization
- Dataset Partitioning
- First Attempts - RNN, LSTM & GRU
- Final Model - Transformer
- Results
- Future Work

Data Processing

INFERNO

- Canto I

Nel mezzo del cammin di nostra vita
mi ritrovai per una selva oscura,
ché la diritta via era smarrita.

Ahi quanto a dir qual era è cosa dura
esta selva selvaggia e aspra e forte
che nel pensier rinova la paura!

=SOCA=

=SOC=

Nel mezzo del cammin di nostra vita
mi ritrovai per una selva oscura,
ché la diritta via era smarrita.

=TER=

Ahi quanto a dir qual era è cosa dura
esta selva selvaggia e aspra e forte
che nel pensier rinova la paura!

=TER=

- It helps to learn text structure
- Tags: “SOCA”, “EOCA”, “SOC”, “EOC”, “TER”

Subword-Level Tokenization

=SOCA=

=SOC=

Nel mezzo del cammin di nostra vita
mi ritrovai per una selva oscura,
ché la diritta via era smarrita.

=TER=

=SOCA=

=SOC=

Nel mez zo del cam min di nos tra vit a
mi rit rov ai per una sel va osc ura ,
ché la dir itt a via era sma rri ta .

=TER=

- It often resembles hyphenation, so it helps to learn hendecasyllabic scheme
- Generation corpus: Divine Comedy; Maximum token length: 3
- 3478 tokens, example: '=TER=', '=SOC=', '=EOC=', '=SOCA=', '=EOCA=', 'to', 'con', 'ta', 'dis', 'te', ...
- Better performance with respect to char level or word level tokenization

Dataset Partitioning

Input

mi ritrovai per una selva oscura,
ché la diritta via era smarrita.
=TER=

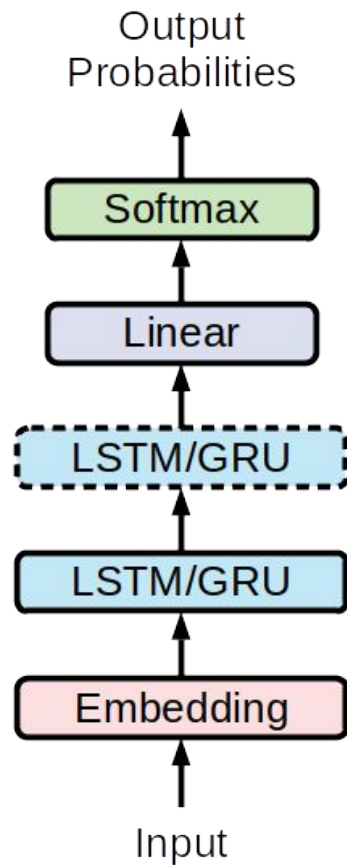
Target

mi ritrovai per una selva oscura,
ché la diritta via era smarrita.
=TER=
Ahi quanto a dir qual era è cosa dura

- *Sequence to Sequence* task
- Input: sample made of three verses
- Target: the input sample plus the following verse
- It helps to learn the rhyming scheme

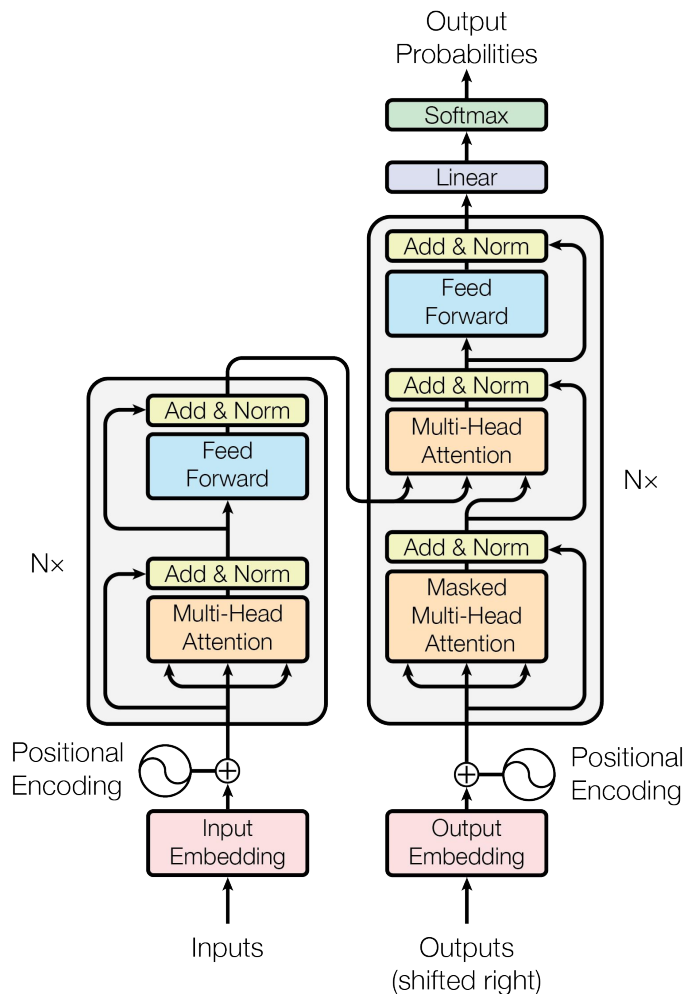
First Attempts - RNN, LSTM & GRU

- Trying de facto standard approaches to text generation
- Bad performance in the generation of a deep canto
- Hyperparameters:
 - Embedding layer dimension
 - Kind of RNN
 - Number of RNN layers



Final Model - Transformer

- Overcome the structural limits of RNNs
- Great ability to generalize in *Neural Machine Translation* tasks
- Hyperparameters:
 - Number of layers for Encoder and Decoder
 - Number of heads for Multi-Head Attention
 - Embedding layers dimension
 - Feed Forward layers dimension



Results

Hendecasyllabicity: 0.904,
Plagiarism: 0.959,
Repetitiveness: 0.998,
Rhymeness: 0.896,
Text Structuredness: 0.996,
Putative Tercets: 62,
Well-Formed Tercets: 62.

We tried several hyperparameters configurations, the best ones produced the metrics above and the deep canto on the right
(*temperature_factor* = 1)

In quel che piace a me l'hanno in ira
fosse in quel che ci ha fatto in ira,
secondo che Dio convien che più spira.

Se venir sù l'allor vi disira",
disse 'l maestro mio senza parte accende,
"quel bene ha mestier di là si tira".

E io l'accorgor; né lo ritene,
ho io d'i pianeti che far l'esser legge,
né ferma qui pena che non m'accende.

E 'l savio come me stesse cade
de la mente lor di quei che 'l ciel chiaro
tutte quante a la quattro strade.

[...]

Results comparison

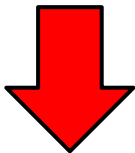
Results of **8** groups are considered

Ngrams plagiarism test	4°
Well formed tercets / Putative tercets	1°
Average structuredness	3°
Average hendecasyllabicness	4°
Average rhymeness	3°

Future Work

From syntactic approach to semantic approach:

The Divine Comedy is plenty of **metaphors** and **allegories**, which bind together images and words. The information is surely denser in words or in a *terzina*, but it is explicitly clearer in the **associated image**.



The neural network should ideally learn to associate words in a given *terzina* and its rhymes to the relative images



Thank you for the attention