

Neural Poetry: Learning to Generate Poems using Syllables

DINFO, University of Florence, Italy DIISM, University of Siena, Italy

Andrea Zugarini, Stefano Melacci, Marco Maggini



SAILab

September 18th, 2019

I C A N N 1 9

Poem Generation

Poem Generation is an instance of Natural Language Generation (NLG).

GOAL: Design an end-to-end **poet-based** poem generator.

ISSUE: Poet's production is rarely enough to train a neural model.

We propose a **general** model to learn poet-based poem generators.

We experimented it in the case of Italian poetry.



We propose a **syllable** language model to catch the style of a poet.

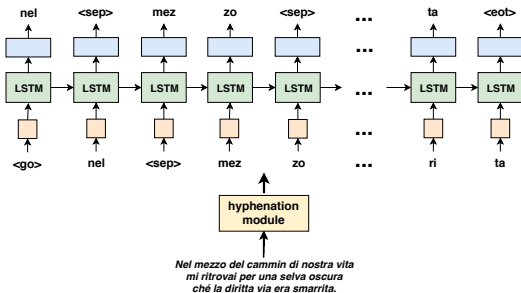
Why syllables?

- ▶ They are the **bricks** to build poems
- ▶ Syllables simplify the application of **transfer learning** from non-poetic languages
- ▶ Vocabulary size is smaller, hence lots of parameters are spared

Syllable Language Model

sy-LM

- ▶ **Input** and **output** text is converted into a sequence of syllables according to the language **hyphenation** rules
- ▶ At each time step t sy-LM outputs a probability distribution $p(x_t|x_1, \dots, x_{t-1})$



Multi-stage Transfer Learning

Training a neural model on single author poems can lead to poor results.

CONSIDERATION: At **syllable** level there are not many differences between poetic and **non-poetic** languages.

Hence, we can **transfer** information by pre-training the model on larger non-poetic corpora.

The model learns in **multiple steps**, so that it can grasp most of the common syntactic and grammatical statistics from multiple data sources.

Once trained, **sy-LM** is exploited to generate new poems, with the following approach:

1. Generate N samples with **Monte Carlo Sampling** from $p(x_t|x_1, \dots, x_{t-1})$
2. Assign a score $R(x)$ to each generated sequence x
3. Select the K sequences with highest score

$R(x)$ is an average of four different functions.

Generation Procedure

R Scoring Criteria

- ▶ STRUCTURE

$$R_1(x) = 1 - \text{abs}(|x| - 3)$$

- ▶ METER

$$R_2(x) = \sum_{v \in x} 1 - (\text{abs}(|v| - 11))$$

- ▶ RHYME

$$R_3(x) = \begin{cases} 1, & \text{if } (v_1, v_3), v_1, v_3 \in x \text{ are in rhyme} \\ -1, & \text{otherwise} \end{cases}$$

- ▶ VOCABULARY

$$R_4(x) = \sum_{w \in x} f_w(x), \quad f_w(x_i) = \begin{cases} a, & \text{if } w \in V \\ -b, & \text{otherwise} \end{cases}$$

We focus on **Dante Alighieri**, the most important Italian Poet.

DATA

- ▶ **DC**: Divine Comedy, 4811 tercets divided in train set (80%), validation set (10%) and test set (10%)
- ▶ **DP**: Other Dante's compositions, some of them are in prose
- ▶ **PAISA'**: a large corpus of contemporary Italian texts

EVALUATION

- ▶ Performances using different training data sources
- ▶ Human assessment of generated tercets from expert and non-expert judges

Experiments

Multi-stage Transfer Learning Results

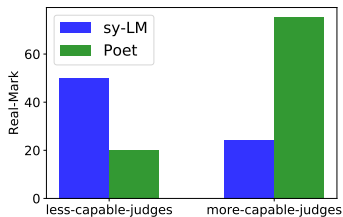
Perplexity on validation and test set, pre-training the model using multiple datasets.

Datasets	Val PPL	Test PPL
DC	12.45	12.39
PAISA' \rightarrow DC	10.83	10.82
DP \rightarrow DC	11.95	11.74
PAISA' \rightarrow DP \rightarrow DC	10.63	10.55

A \rightarrow B means that we train on A first, and then we train on B.

ANNOTATORS' TASK: decide whether a tercet was real or not.

Generator	Real-Mark
sy-LM	28%
Poet	64%



Experiments

Expert Judges

Four academic Dante's experts were asked to judge different properties of our generated tercets.

	Readability	Emotion	Meter	Rhyme	Style
Judge 1	1.57	1.21	1.57	3.36	2.29
Judge 2	1.64	1.45	1.73	3.00	2.27
Judge 3	2.83	2.33	2.00	4.17	2.92
Judge 4	2.17	2.00	2.33	2.92	2.50
Average	2.04	1.73	1.90	3.37	2.49
Poet (Average)	4.34	3.87	4.45	4.50	4.34

Each expert evaluated 20 tercets, 10 generated and 10 real.

*e tenendo con li occhi e nel mondo
che sotto regal facevan mi novo
che 'l s'apparve un dell'altro fondo*

*in questo imaginar lo 'ntelletto
vive sotto 'l mondo che sia fatto moto
e per accorger palude è dritto stretto*

*per lo mondo che se ben mi trovi
con mia vista con acute parole
e s'altri dicer fori come novi*

*non pur rimosso pome dal sospetto
che 'l litigamento mia come si lece
che per ammirazion di dio subietto*

Project reference sailab.diism.unisi.it/poem-gen/

Conclusions and Future Works

- ▶ We presented a general poet-based poem generator based on syllables learning from multiple corpus
- ▶ Generated tercets are considered to be real by evaluators with humanistic background roughly half of the times of Dante's verses
- ▶ Expert judges appreciated *rhyme* and *style*, while *meter* and *emotion* are not captured yet.

Next Steps

- ▶ Use the scoring criteria in a Reinforcement Learning strategy
- ▶ Enrich the model with additional features and information about the author and its style

Thank You !!!

Project reference
`sailab.diism.unisi.it/poem-gen/`