

Poems generation

Karolina Bogacka & Marek Brynda



Actors of the play

- GPT2 - papuGaPT2
<https://huggingface.co/flax-community/papuGaPT2>
- Datasets:
 - Pan Tadeusz
 - Poems by Kochanowski and Tuwim

No training

Input:

Litwo Ojczyzno moja ty jesteś jak zdrowie

Output:

Litwo Ojczyzno moja ty jesteś jak zdrowie

Do tej dziewczyny, w której zakopałeś swoje buty nie ma nic, czego by nie można byłoby nazwać "pozornie niewinną istotą". Nie jest nią ani sam mężczyzna, ani jego towarzyszka. Nie jesteś tym, który ma prawo wiedzieć, czego chce, a co jest bezcelowe. Nie jest nią również ten, kto nie chce wiedzieć tego, co mu najbardziej leży na sercu. Kobieta może cię nie lubić. Jej charakter będzie cię za

No training - context includes rhymes

Input:

Nad rzeczką opodal krzaczka,
Mieszkała kaczka-dziwaczka,
Lecz zamiast trzymać się rzeczki,
Robiła piesze wycieczki.

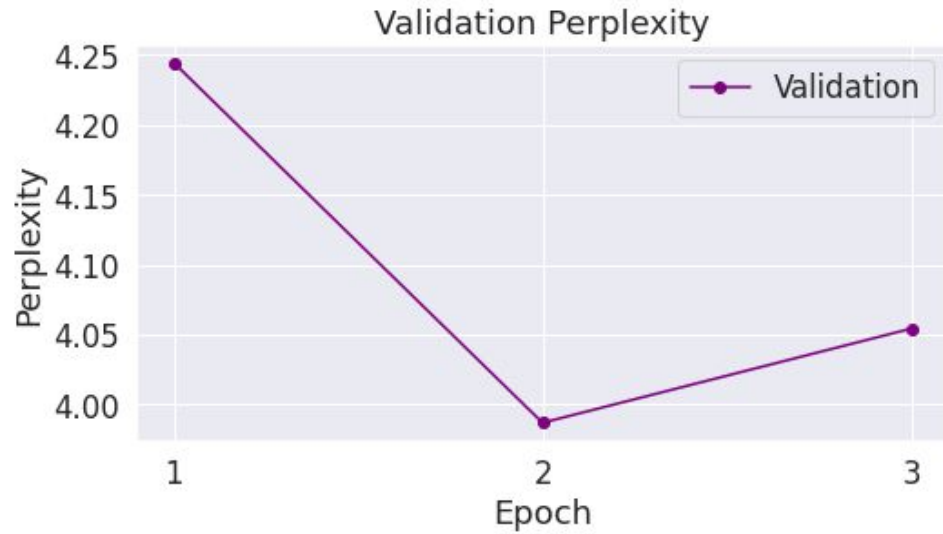
Output:

Nad rzeczką opodal krzaczek
Mieszkała kaczka-dziwaczka,
Lecz zamiast trzymać się rzeczki
Robiła piesze wycieczki.

Przebiegała tędy i pod górkę, przez las. Zwała się Roszowica. Miała na imię Lechonia. Lechonia była nazywana od rzeki. Była najstawniejszą kobietą w całym wszechświecie. Jej imię oznaczało "zakochana"
To była historia o tym, że nie można

Fine-tuning





Stopping criterion

perplexity of validation set is increasing

Pan Tadeusz



Training using Pan Tadeusz

Input:

Litwo Ojczyzno moja ty jesteś jak zdrowie

Output:

Litwo Ojczyzno moja ty jesteś jak zdrowie
Jako święty Jan Kapitolu, którego pamięć,
Miejsca, czas i myśli, które niegdyś wypełniały
I czczone były w Rzymie jako święty Jan Trzeci;
Jan Trzeci, jako najwyższe męstwo i zasługi
Rad był dumny i szanowany na Litwie za czasów Austryjackich.
Ale młodzież, słuchając i poddając się, na koniec wyrzekła,
Że Pan Bóg za grzechy nie przebaczył.

Training using Pan Tadeusz

Hrabia z przymrużonym okiem
Stanął i stał z przymrużonym okiem,
I nie poznał ich dotąd w swoim życiu.
I rzekł: «Tak widzimy, że...; i ja to znam!»
I dalej nie wiedział, co ma odpowiedzieć,
I zrazu nie śmiał otwarcie odpowiedzieć,
Już potem — lecz sam w gniewie krzyknął: «Jeszcze nie wiedzą,
Czy ja się w swoim własnym życiu od nich zachowam,
Na wypadek, gdyby mnie ksiądz kapłanem...
Nie było w tym nic złego!...»

Training using Pan Tadeusz

Mrok zaścianku przygasnął, jakby od świata zstąpił.
Teraz, za wcześniej, za wcześniej i za wcześniej,
Już późno widać, już widać: i tak świat ginie
W mgnieniu oka. Już Zosię Zosię przelekła,
Aż wreszcie z rozpacz i przerażenia,
Już zaczęła szemrać i coraz z niej walkę wyręczać.
Boi się, że gdyby się rozstała pieśń ludowa,
Musiaby ją wkrótce zabić, niż myślisz.
Mrok się zawalił i cała szlachta zamarła przed domem;
Stąd też mrok już dawno zgast; lecz wszyscy milczą.
Wszyscy milczą, każdy na własną modłę:
Wszyscy rychło z żalu i rozpacz, ci, co milczą,
Podburzają; sami się przeklinają z domu wznosząc;
Więc za czym płaczą litościwi, na których się trwożą?
Któż ich broni zmówić poranka, kto przy świecy
Nie odmawia pacierza, pacierza...
Tylko kto wie, w jakim gniewie powstanie?
Kto w gniewie się gniewa, kto w gniewie nie umie?
Kto nie podda się gniewowi? Lub za czym się zapali?
Kto wie, może trzeba do niego postać urzędnika?
Wzywamy: jeśli ten, co w domu? jeśli w Litwie? cóż na to?

Wojski, jak pan — rzekł Rejent — gdy im czas zasępi
Wszakże pan do mnie na sąd go po tylu latach
Pochwalisz? a to teraz powiem. Wdzięczna rzecz;
Obyś się stał po zdrowie do mego honoru,
Tylko mnie na to czas znajdziesz. Od czasów sejmików
I sejmików mych do woli gadaliśmy
I do mojego dawnego nazwiska się nazywałem Rejent,
Który na sejmikowych zjazdach prawych przed sejmikiem
Wstrzymał wszystkie sejmiki, wszystkie sejmiki,
I na sejmikowych zjazdach jaśniejących rad byłem rad
I do moich dóbr nie należę;
Jakże więc, kiedy staję przed trybunałem wojewódzkim,
Ja muszę głosować, a potem, kiedy staję przed trybunałem
Zakonu? *!*!* — *!*!*...!*...
A nie... ja muszę brać, ja muszę dostać ingrediencję*.
Gdy z nudy i hałasu, słuchając rad Wojskiego,
Zapomniałeś, lub nie pojmałeś w gniewie twojego ducha;
Dałeś się porwać wstawom i w układy,
A teraz do skutku za Panem Rejentem wolę;
Bądź zdrowa! Robak, jak to szlachta — szlachta do skutku.
Już dawno temu mnie opuścił talent,
Uczułeś, że ci to pany mają w sobie uroku?

Poems



Training using poems

Input:

Litwo Ojczyzno moja ty jesteś jak zdrowie

Output:

Litwo Ojczyzno moja ty jesteś jak zdrowie

Imię twe kochane!

Imię twe wielkie,

Twoje imię sławne,

Imię waleczne,

Swoje prawdziwe,

Święto, które wszystko moja ręka

Jest w tobie znamieniem

Imię święte,

Imię, które stało się tobie

Oliwierowi,

Imię twoje ukochane,

Imię twe ukochane!

Imię twe kochane!

Training using poems - context includes rhymes

Input:

Nad rzeczką opodal krzaczka,
Mieszkała kaczka-dziwaczka,
Lecz zamiast trzymać się rzeczki,
Robiła piesze wycieczki.

Output:

Nad rzeczką opodal krzaczka,
Mieszkała kaczka-dziwaczka,
Lecz zamiast trzymać się rzeczki,
Robiła piesze wycieczki.
Tak po lesie, jak i po sadzie
Baby krzakiem sadził,
I dziwił się pies, i dziwił się
Że kaczka
Chodzi, by ją mieć ze sobą.

Chocia nie dziwił się pies, bo w borze ją lubi,
Że tamtędy chodzi, a ten większy,
By ją mieć ze sobą.

Ale jak ją miał położyć,
Kiedy był w żmijach,
Śmiesznie płakał.

Training using poems

1.

Owa, o śmierci, śmierć moja!
Tyś ta, kto, mię kochasz, jesteś, jedyna nadziejo!
Żeś mię tak wiele raniła, nie mogłaś więcej.
Nie mogłaś ani więcej żyć wiecznie,
Ale mię z nadzieją i z żalem ranisz.
Teraz w oczu twych widzę zwierciadło,
Abych więcej widział i umiał, niżli śmierć tak wiele sobie
A propos roić mógł, smutniej to widzę.

2.

Zapomnisz o tej chwili,
Tak jakoś dzień po chwili o nim opowiadasz.

3.

Ach, żale na ziemi!
Ach, żale na ziemi!
Ale kto się chce ostać,
Nie ubędzie; nie ubędzie w miłości,
Bodaj się uchowa, tylko dostanie,
Ale już będzie żałować,
Ze swej strony, ale się uchowa.

Looks nice but...

No rhymes

Possible approaches:

- **Generate line
& next line - 1 word**
**Then try to generate multiple list of
tokens ending with \n and find the
one with the best ending syllables**
- Generate line
Try to generate multiple next lines
and find the one with the best ending
syllables

Force rhymes - end of 2nd sentence approach

1.



2.



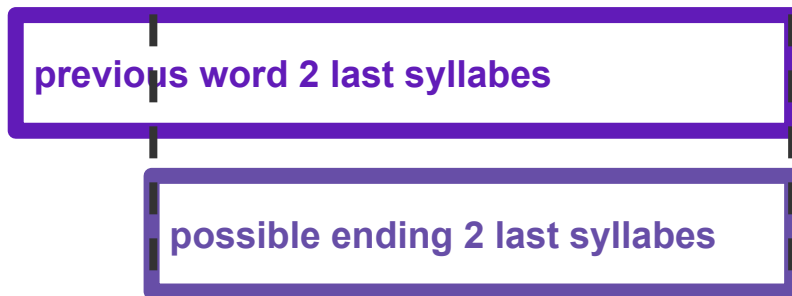
x N

3.

Find the best fitting last word before \n (comparison of last 2 syllables)

Force rhymes - end of 2nd sentence approach

- 4.** For each possible ending take last word, take last 2 syllables
If syllables from the word from the previous line are longer, then cut them.
Otherwise cut the possible ending syllables to match the length,



then compute Levenshtein distance and divide by the common length (length of originally shorter string)

Force rhymes - end of 2nd sentence approach

Input:

Litwo Ojczyzno moja ty jesteś jak zdrowie

Output:

Litwo Ojczyzno moja ty jesteś jak zdrowie

Na drugie więc nie masz już polegający na zdrowiu

Na drugie więc nie masz już powodu,

Aby człowiek pujemy się na gruzach Babilonu

Na drugie więc nie masz już żadnego skutku.

I nie masz już ani jednego, ani polegającą skutku,

O pomocy nie masz już nic.

Force rhymes - end of 2nd sentence approach

1.
Już mi radzili
Słowa nie ianthi zybili,
A ci, co nie wierzyli,
Przypuśćmy, że nie ramionami tworzyli
Na miłę; ci, co nie wierzyli,
Przypuśćmy, że nie chcieli użytek czynić.

2.
Bartoszu,
Okryj się, dziecko, interpretatorem swej duszy
A na oczy i na patrzatki nadziejowe
Unieś do nieba gwiazdę z Centre furoreowymi
warekowymi kominami,
ramionami, kominami,
Wzgórę kominami

3.
Królewno piękna, z prośbą, o zdrowie,
Królu wieczny, któremu służą zmatowie
Dobrodziejstwa, przebacz me nieprawości!
O zdrowie, o życie sprzedawany Bogu i zamierzonych trudności,
Które dziś w sercu twym mieszka,
Daj, abych ja i ona zdrowieła w ciele.

Summary

- Applying only learning using poems
- not enough to teach the model rhymes
 - Usage of archaic language could be seen as disadvantage - or advantage
- the strange looking words are not standing out so much
 - Applying simple rule helped find structures similar to rhymes - but it is kind of slow and not the perfect approach
 - Using stopping criterion on EOS token or dot after reaching 75% of expected tokens
-

Failed approaches

- We tried T5 model with “forced” tokens - unfortunately it was hard to force it to use them at the end of next line
- Idea of using Masked BERT - not enough context to generate whole new line, problems with generating beginning of the word knowing the last syllable

Feedback

- Having more time and resources we would try bigger model GPT-XL or MT5-XXL
 - Better prepare input dataset - understanding better models we could focus on getting only a specific type of rhymes or use modern language...
 - Could try approach with custom tokens focusing on rhymes
-