Transformer

# DATASET: 30 MB of Polish novels

## Around 220 000 unique words among around 5 000 000 words

### Inicial attempts – parameters and architecture

**Samples**: Sentences & padding

**Punctuation**: fully removed

**Restricted number of words**: -

**Model**: Small model with TokenAndPositionEmbedding layer and one TransformerDecoder

* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 5, **Adam learning rate**: 0.001 🡪generate nans as weights
* **Batch**: 16, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 2, **Adam learning rate**: 0.001 🡪generate ResourceExhaustedError
* **Batch**: 8, **Embedding dimension:** 64 , **Number of heads:** 4, **Epochs**: 2, **Adam learning rate**: 0.001 🡪generate ResourceExhaustedError
* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 2, **Adam learning rate**: 0.0003 🡪generate nans as weights
* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 1, **Adam learning rate**: 0.0001 🡪generate nans as weights

**Others**: sentences longer than 40 were removed, samples shuffled, “Sequential type of input is incorrect for parallel processing model” – resigned from using generator, replaced with Tensorflow Dataset

The problem discovered is nans as weights. Looks like a gradient explosion in a tiny model.

*Saving problem -* after changing the layer names for unique ones, and weights names for unique ones, the optimizer weights names for the unique ones - the model can be finally fully saved.

**Model**: Small model with TokenAndPositionEmbedding layer and two TransformerDecoders

* **Batch**: 8, **Embedding dimension:** 8 , **Number of heads:** 2, **Epochs**: 5, **Adam learning rate**: 0.001 🡪 generate nans as weights

### Attempt with restricted words number

**Samples**: Sentences & padding

**Punctuation**: fully removed

**Restricted number of words**: 100 000

**Model**: Small model with TokenAndPositionEmbedding layer and one TransformerDecoder

* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 3, **Adam learning rate**: 0.0001

**Others**: sentences longer than 40 were removed, samples shuffled

**Results**: weights ok

za górami za lasami gdy to gdy gdy a a po po ale kiedy kiedy to kiedy gdy tak

z czy co niego się on gdy to w z na jak to ale do się siebie już że od i nemo

+

**Epochs**: 3

**Results**: more words appeared (besides stopwords)

Za górami za lasami i to a w na po to tak i gdy z gdy po ja ja lecz ale nimi

siebie o potem tym za po tego prostu ci co że tak pomocą dnia że ja i ma mnie

był nie nie to lecz w a i a na nie z na i nie i to zaś do jak mógł co gdy za

mógł świecie będzie znaczy tego go mi to z człowieka nie tych chcę tak kiedy

kiedy i co na to na z w to – i na to z – ale co w zaś się wszystko nim był z był

co do z w tu to głowy tym wszystko kraju w nie gdy co to ale z po a gdy i kiedy

nie – na lecz na ale co tej król się tej za się mógł niego nią to zaś samej to

ale że przyczyny tak woli a lecz kiedy i – – tak tak na z tak tak a nie z tak –

na i on nie ale mi koniec nie teraz sobą i sam ma że był że się ja już tego

+

**Epochs**: 4

**Results**: even better

Za górami za lasami gdy a gdy ale nie ale a co tak ale lecz gdy w tak na ale ja

chwilę to nie to ma na długo tak ręku nich się za na nich to za każdym i mną

słowa jeśli a kiedy kiedy w a tak z tak gdy z i ale nie ale był – miejscu i pan

go domu tak tak początku z i rzekł tego to w jest na co nich świecie może gdy i

– nie gdy a ale ale gdy po – – a a nie ja co chwila z już wyrzekł mu co już za

nie było za sobą ale się nic na na za niego sobą nie to z w z co gdy co i a lecz

na w gdy w po co koniec na mogę tego jego pan to tak życiu nich by czele się

andrzej że i tych nie tak można to po w ale gdy i na to to z nie tak co po ja

ale gdy kilku jak na ten to to rzekłszy wiedział to nie się to człowiek na nie

ma by koniec nie domu

+

**Epochs**: 5

**Results**: no relevant improvement, even worse

Za górami za lasami po ja gdy a tak z a ale gdy a na co – a i ja tak późno nią

prostu to było z go dworze się po było to to wolna go tym już do z warszawy w

nie po gdy po z i gdy gdy a a ale tak to – a gdy była tej tem tym tym ja się nie

mi zwrócił się mierze sposobem nie wiem mu w był na pewny to a kiedy w kiedy tak

po – – gdy i co gdy tak i nie tak końcu świecie na mnie zaś czym co w w to długo

o tak był niej było niej na tylko niego gdyby na nie nie a i ja w na w co po i

po – ale nie w po był mogąc może też tem ty ja co może tym ich się się to nie

jak być mi nie w – i ja nie kiedy na w nie tak tak i to nie po po miał z mógł

odparł mam zaś istocie nie się wiem to nimi julian jak będzie to na nie za

wiedział

# DATASET: 10 MB of Polish fairytales and stories

## Less than 92 000 unique words among around 1 300 000 words

### Inicial attempts

**Samples**: Sentences & padding

**Punctuation**: fully removed

**Restricted number of words**: -

**Model**: Small model with TokenAndPositionEmbedding layer and one TransformerDecoder

* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 5, **Adam learning rate**: 0.0001 🡪 generate nans as weights
* **Batch**: 4, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 5, **Adam learning rate**: 0.0001 🡪 generate nans as weights
* **Batch**: 4, **Embedding dimension:** 32 , **Number of heads:** 1, **Epochs**: 5, **Adam learning rate**: 0.0001 🡪 generate nans as weights

**Others**: sentences longer than 32 were removed, samples shuffled

**Results**: generate nans as weights, maybe because we have three times less sentences than in previous attempts (285 068 🡪 98 829).

### Attempt with restricted words number

**Samples**: Sentences & padding

**Punctuation**: fully removed

**Restricted number of words**: 60 000

**Model**: Small model with TokenAndPositionEmbedding layer and one TransformerDecoder

* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 4, **Epochs**: 5, **Adam learning rate**: 0.0001

**Others**: sentences longer than 32 were removed, samples shuffled

**Results**: weights ok

Za górami za lasami a kapitan z co nie z to w na o co kapitan ale nami sobą

nawet tej chciał co ten na to w kilku chwila to nocy się kapitan o gdy po czy

ale ale czy w tak kapitan ale z o tylko w to chwili nie może jak to jego jego po

nie było nas jestem kapitan po co z gdy na nie co co w nie ale ale tak już o

drodze powodu chwilę się ciągu kapitan to co ten pod nie nam wodą statku po o po

ale w nie co w a czy w kapitan o miał która kilku prostu tej tu już górę tej do

że jest i góry w kapitanie a a gdy to to kapitan nie ale a tak tak na na do nocy

co go jeszcze widziałem on o nim bo już w nie ja końcu są

+

**Epochs**: 10

**Results**: better

Za górami za lasami zresztą po z ale w to kapitan tak a z co a czy każdym chwilę

nie wody odległości nie to w jest bo przez miał tym nigdy słuszność ich zresztą

zresztą kapitan co gdy nie co tak na ale nautilus nautilus kapitan że panie jak

farragut w tylko powierzchnię tak nemo profesorze kapitan tylko się jak przez go

po ale kapitan w nie co w nautilus z tak z ale conseil w kilka upływie żeby ma

każdym tego nas który kroków by rokiem do przy na tym kapitan i i na kapitan a

co i tak nautilus nautilus to tak tych jego kazał nautilus niż w łatwo był

dobrze statek zdawał której od odpowiedział był mowgli kapitan tak ale w

nautilus na na tak nautilus zresztą gdy a nie nas tak i po jest powierzchni w

ned był długo pewnym morza landa tak się bardzo

+

**Epochs**: 10

**Results**: no improvement, even worse

Za górami za lasami po nie tak na kapitan tak gdy z nie a ale nie na nami to tym

że nie jednak widziałem co jego tylko tak kapitan pan nie moi mogłem na nautilus

w na kapitan po kapitan zresztą kapitan i nautilus co a pan towarzysze pokład

tej nemo nie by zbliżył gdy moi rzeczy ma na także słuszność mówił nautilus gdy

to i z i nautilus po czy nie tak ale nie to ani zdawał już jego na mi że miał

tego po to są z jest powodu w ale nie zresztą tak zresztą to zresztą w nautilus

na po nautilus za jak wodzie się że w tych powierzchni w w już której których

tej było samej po to i po w zresztą ale kapitan to kapitan to z na się dla

trzech ja ciągu pan na że powierzchnię których do się są te od dzieci

+

**Epochs**: 15

**Results**: no significant improvement

Za górami za lasami po w i po tak na kapitan i kapitan i nautilus czy nie mną

chwilę chwili pod bardzo kapitan wziął z się potem morze czy sobą zaczął to

wołać gdy kapitan zresztą w z po to tak tak a po ale zresztą a pan mnie w

których co że tym na chcesz każdym by względem go ich więcej kapitan zresztą gdy

z zresztą gdy nautilus po to gdy nie nie nautilus jeżeli było wziął w jest stał

jeszcze ma na tak takim przy słuszność mi ich głowę kapitan w w z i tak w

nautilus zresztą na na kapitan kapitan i i nemo każdym z tym panie powierzchnię

kapitana w razie miejscu morza tych jego zwierząt gdy nie na a i i nautilus ale

gdy i nautilus z co po a nagle tym się był w stanął w może którzy bardzo nagle

na nie statek

+

**Epochs**: 15

**Results**: even worse

Za górami za lasami na zresztą nautilus w i zresztą i z w i z z w mało nimi jego

zbliżył że ówdzie ten tej każdym to mówił luźne ziemi na te platformę po na na

po nie tak się tak zresztą tak tak kapitan tak nie same drugiej godzinę jest jak

co jak na wspomnienia potem mi mi ich je wyrazu gdy ale i z się po nie nie tak

na tak w i ją pod tylko w że można się łatwo nie wieczór tym było było zapadł na

ojciec i na nie to się i nautilus tak po nie zresztą nie i na to nie miał że

zbliżył południu od że że co tylko nas z w powodu lecz ale kapitan to i a gdy

tak ale ale zresztą to to i każdym nautilus tych na w ned co się samym którzy

ciągu do o chwili sto

# DATASET: 5 MB of Polish children stories with simple vocabulary

## Less than 73 000 unique words among around 860 000 words

### Attempt with restricted words number (just slightly)

**Samples**: Sentences & padding

**Punctuation**: fully removed

**Restricted number of words**: 70 000

**Model**: Small model with TokenAndPositionEmbedding layer and one TransformerDecoder

* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 8, **Epochs**: 10, **Adam learning rate**: 0.0001

**Others**: sentences longer than 35 were removed, samples shuffled, testing for more heads, but using simpler vocabulary, that is almost fully included in the dictionary this time, and smaller dataset

**Results**: no strange words

Za górami za lasami o następnie nie co w w ale ale po o na pewnego kiedy po gdy

mógł aż w ci górę tak ile pewnego śmierci się jego w mi dla miejsce nie może i w

kiedy i w pewnego gdy kiedy o jestem i kiedy na w ale nie już jej poszedł dawna

ją co mu górę jestem się i tu na w nie wodzie o gdy na po po o i kiedy w pewnego

w gdy król ale to domu a z kilku tym król jednak się teraz z powrotem celu tak

co do wielkie mnie to na ale to gdy i gdy tak król w o nie na kiedy i lasu po

koniec jest dlatego mi nim mógł znów i nie żeby za ze było chwilę bardziej po

pewnego nie jestem kiedy kiedy i kiedy kiedy jestem na z o gdy tak dla i jednego

i już ją i nich zaczął swoich i nigdy zapytał się nie głosem mam

+

**Epochs**: 45

**Results**: lots of stopwords, maybe some improvement

Za górami za lasami ale a nie to o w z a na w gdy z to gdy ale za i na ty każdym

ja pałacu jego nagle wiele ziemię słowem w i głos jej młodzieńca ale nie a w po

pewnego to król nie to ale to król na gdy mój że mogła mgnieniu ranka posłał za

już widok boże się lecz nią a za i rękę i to i ale ale król nie po kiedy po po

nie ale to następnie mąż tak jest król kazał pewnym prostu jest znaczy nie że ci

plecami że ma nie słowa po z i w nie ale w i w na o król król pewnego nie w ale

niego tym w nie imię i czasu świecie noc którym i z i powiedział uśmiechu z i w

w po z gdy kiedy w pewnego w kiedy na a i udał nie rzekł mgnieniu pewnością ją

gdy z ona bo w że młodzieniec ma tym się samym

+

**Epochs**: 25

**Results**:

Za górami za lasami a i ale gdy w gdy to to to kiedy nie a to to ale bardzo nad

z się mu za zaś następnie samo daleko nim było miał tak się je podział i ale

król tak gdy po a ale a i o król z gdy nie już się po łatwo drodze w rzekł

poszedł go w drodze ale ojciec w przed i oczyma na był w na to pewnego a kiedy

ale w pewnego po tak tak a tobą przez szczęśliwy pewno czasu to lesie czym łatwo

ale król ale a sobie posłał na syna gdy nie kiedy król ale król z w na po

pewnego ale pewnego w gdy zaledwie lecz miała i był domu mieście król oddali

trzy z jeszcze i się nią z głowę a ale nie z o a z nie gdy nie nie z ale gdy

kiedy i i na nich to było miała tego młodzieniec jeszcze nią w to miał nie samo

mam

# DATASET: 2.3 MB Miniset of the most famous short children fairytales

## Around 40 000 unique words among around 350 000 words

### Attempt with all words included

**Samples**: Sentences & padding

**Punctuation**: fully removed

**Restricted number of words**: 38 377

**Model**: Small model with TokenAndPositionEmbedding layer and one TransformerDecoder

* **Batch**: 8, **Embedding dimension:** 32 , **Number of heads:** 8, **Epochs**: 50, **Adam learning rate**: 0.0001

**Others**: sentences longer than 36 were removed, samples shuffled, testing for more heads, but using simple vocabulary, that is fully included in the small dictionary, smaller dataset whereas more epochs

**Results**: the best so far, fairytale vocabulary

Za górami za lasami król król nie nagle a po nie i następnie w nie po

młodzieniec i ale chwilę sobą powierzył będziesz jednak jestem następnie ma

który ginę do miała córką przed się jeszcze słońcem człowieka kiedy to kiedy

potem po kandyd a nie następnie król na młodzieniec kiedy i a jej ani nagle

ujrzał skończonej teraz po stole już złotopiórcia śladu las przecież chociaż i

otoczony na odjechał to nie wieczorem młodzieniec król król w w nie dnia o

młodzieniec klara ale gdy opowiedział chwilę że dotarł przyjął drogę wolno tym

zaczęła tymczasem choć przez żeby co graniem chwilę dzień zatrzymywały kiedy

klara gdy a kiedy nie gdy w kiedy dnia ale w w nagle następnie dziewicę cały

znów ujrzał się spostrzegł za wciąż mgnieniu książęta był ogniste więc pilnowany

dobrych domy w stopy rozdział król pewnego kiedy ale nie na nie po dnia pewnego

po na z ale które końcu xiii młodzieniec król to cichutku ranka widok odgadła

jednak który ty szeroką nosi z śmiertelnym powrotem

+

**Epochs**: 30

**Results**: more stopwords?

Za górami za lasami po kiedy gdy król i a o na i młodzieniec kiedy a król z na

chwilę duże rowach nie nie mało zwieszała znów po pogawędki a zjem z dotarł

rozpostarte tak tam drzesz kiedy gdy gdy po na gdy o i na dnia następnego

młodzieniec po i na za do byli do klombach tym łące księcia chwili poddasze

tańczącego nich nie klaro także jak rzekł zaczarowane kiedy na to a tak kiedy

pewnego gdy po a kiedy i pewnego gdy król o w nagle że się przyszli czym

doralice wieczoru spostrzegł dalszą z w kawalerem to niej przywitała pazurami a

gdy więc wieczorem a w o kiedy król nie ale na o a król był jej wiatr płacili w

tkaczach udał już swoim myszy mowę… dla nie za polnej gościa mną serdecznik na

klara to w król w w kiedy a i po król król nie nie jest gdyż podwórzu nie kazał

środku młodzieniec czym nie mogła nie chce tryskała księżniczce udać je czytać

sprzedać

+

**Epochs**: 50

**Results**: better story?, but “i”repetitions

Za górami za lasami po to i i i ale o a po na kiedy w to o po mąż pierwszym jej

uciekł z pień cichutku ujrzała przed horyzont gdyż ile twarzy i i by kacze

spalono młodzieniec nie nie wieczorem o gdy król pewnego na kiedy gdy na w król

a opowiedział podwórze rzucił będziesz litość posłał myśl tylko kołysce macocha

i w pierścień z jest swoim mojego przyjacielem książę więc w to tak to nie z

złotopiórcia kiedy a kiedy w na ale bardzo państwa nakazał końcu już troszczyła

piękna co końcu nie mojego na mi to chce to go wierzyć a na pewnego wieczorem

nie a gdy a król to a kiedy kiedy w następnie znowu nie gdy wieczora był król

pojął gdy na torbę ma tropił i kukułka kotarami córeczkę zetną skrzynkę klara

ale i nie a i to po nie nie o po a a po który przez po z jednak tylko wolno

świcie co dotarciu niego rąk co młodzieniec przy oddali zrobił dzieci