GPTSANプロジェクトの実施結果

異能vationプログラム破壊的挑戦部門
【GPT-3相当の大規模言語モデルの日本語版学習済みモデル作成】
で作成を目指している大規模言語モデルの開発プロジェクトの実施結果について

開発の実施内容

実施した開発内容

→Switch Transformerモデル設計。

→モデル並列による巨大パラメーターモデルの学習プログラム作成。

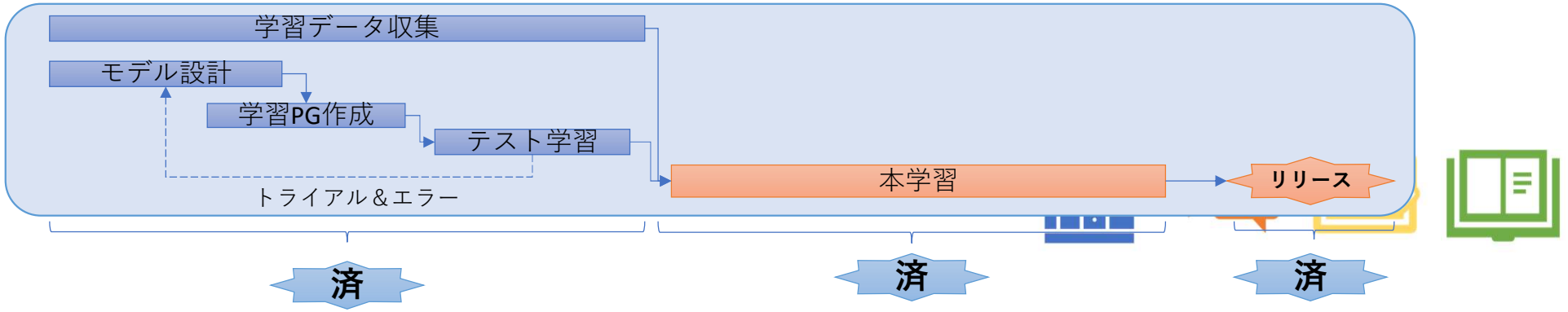
→TPUによる学習が出来るように、学習プログラムを改良。

→専用の日本語エンコードアルゴリズム開発（36Kトークン）。

→学習用データの収集（500GiB～）。

→本学習（500GiB～）。

開発フローと現在の地点

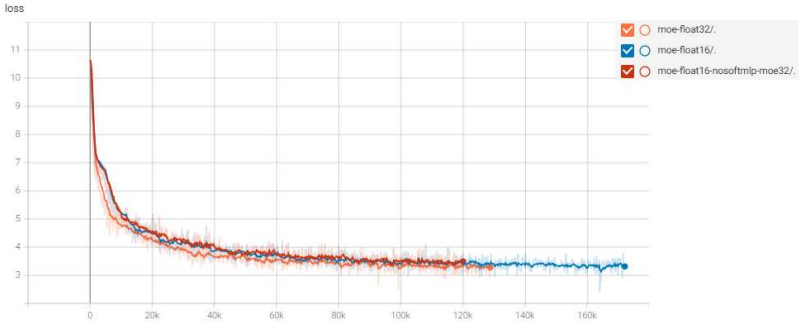


既存モデル（GPT-2日本語版）との違い

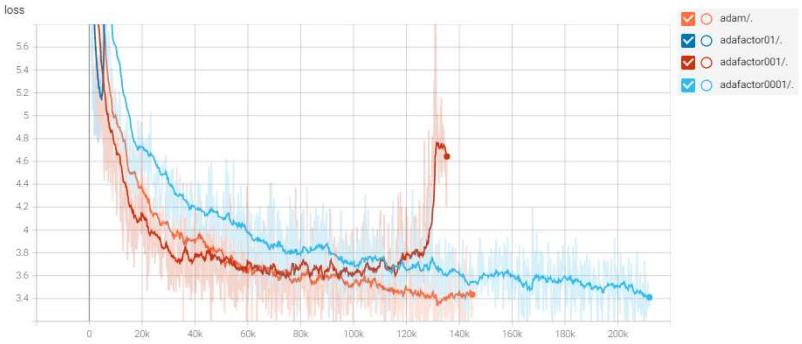
モデル	項目	内容
GPT-2日本語版 （既存モデル）	モデル形式	Transformerモデル
	学習データ量	20GB程度
	パラメーター数	324,426,752 (3億)
	学習手法	データ並列
GPTSAN テスト学習	モデル形式	Switch Transformerモデル
	学習データ量	10GB程度
	パラメーター数	2,776,248,480 (28億)
	学習手法	データ並列 ＋ モデル並列
GPTSAN 本学習	モデル形式	Switch Transformerモデル
	学習データ量	500GiB～
	パラメーター数	2,776,248,480 (28億)
	学習手法	データ並列 ＋ モデル並列

追加テスト学習ーモデル最適化

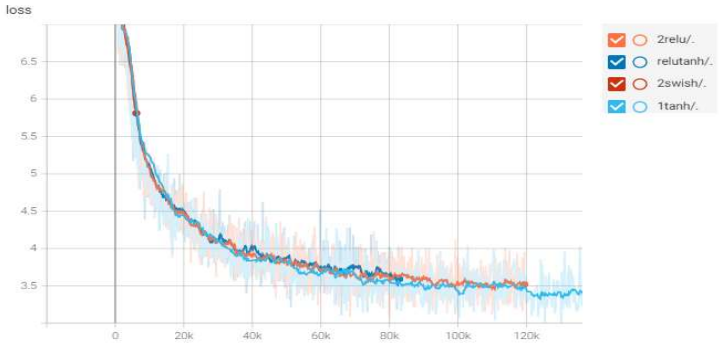
Mixed Precision Training (16bit-float)



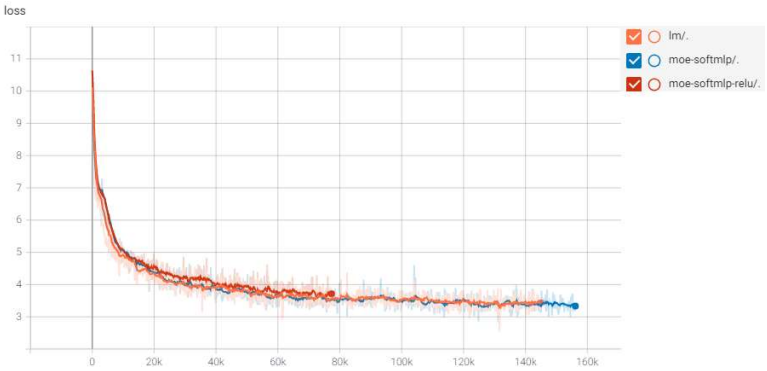
Optimizer Valiation (Adam vs Adafactor)



Activation on soltmlp



Soltmlp (Skip Connection on MLP)



学習時の損失

GPTSAN-2.8B（本学習）の学習曲線

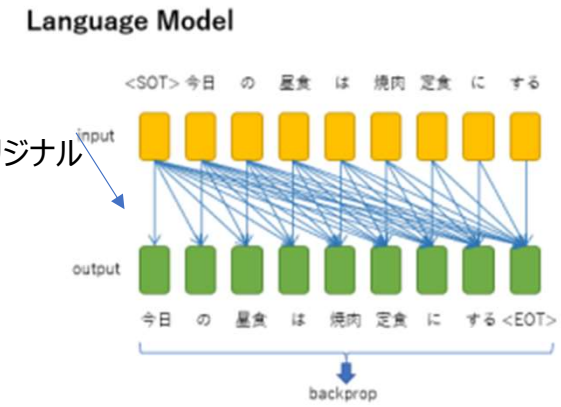
目標値の**2.3**前後に到達

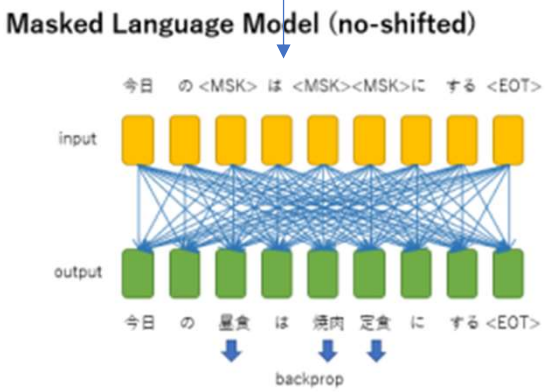
GPTSANオリジナルポイント

Hybridな言語モデルを作成

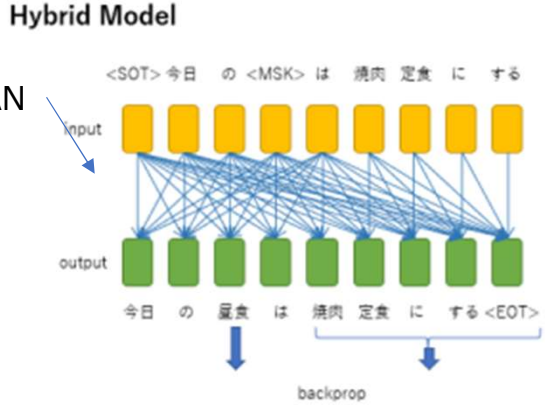
BERT

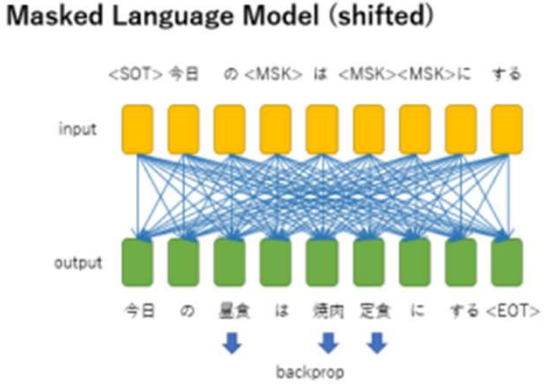
GPT-2/
GPT3オリジナル



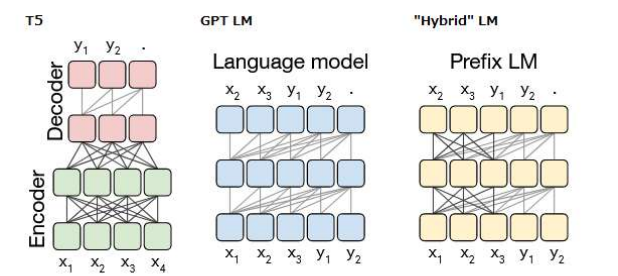


GPTSAN





参考：T5論文（google）

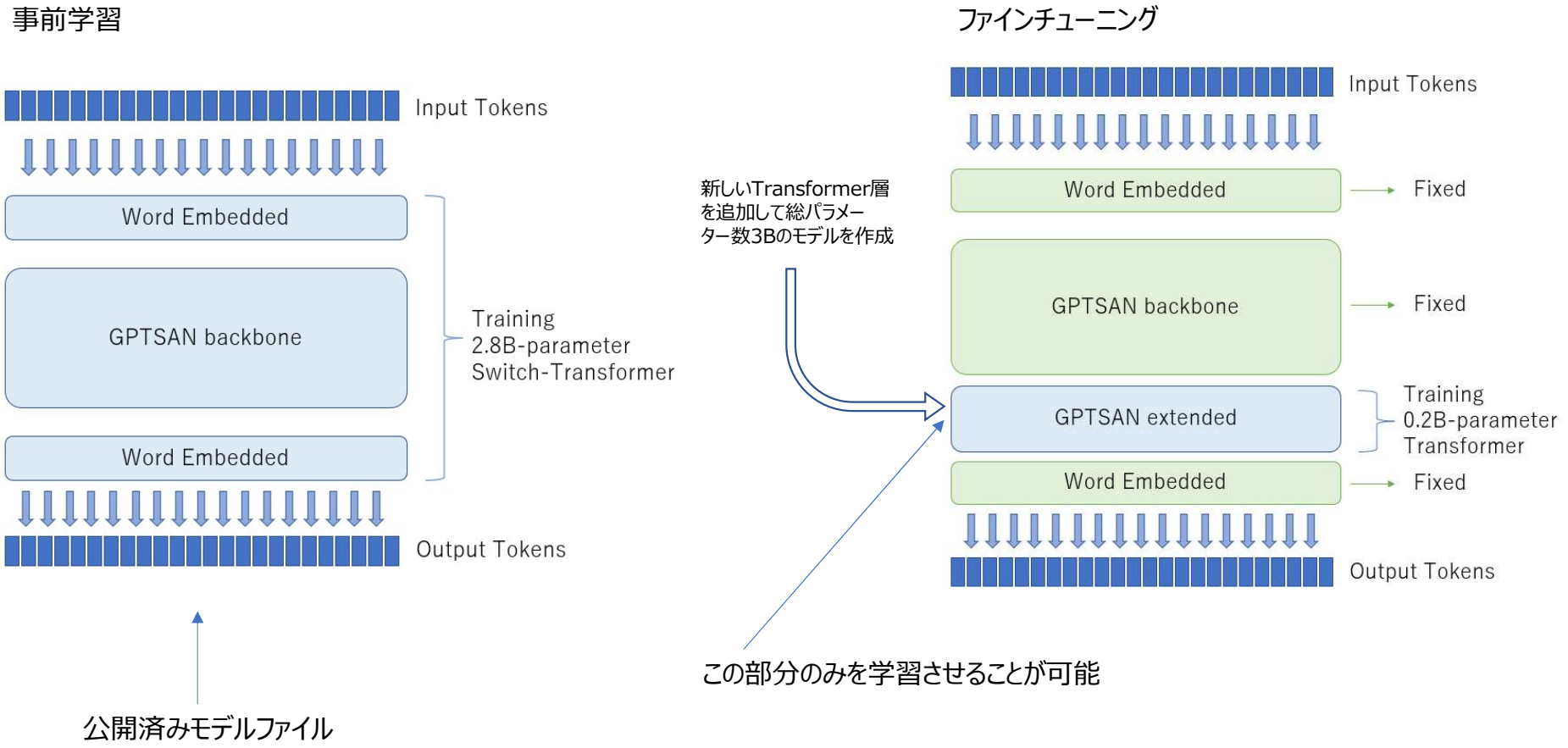


Hybrid相当

T5の論文によると、ObjectiveをLanguage Model
とした時の最も良いモデルはPrefix LM
文章の続きを生成したり、抽象型要約に向いている

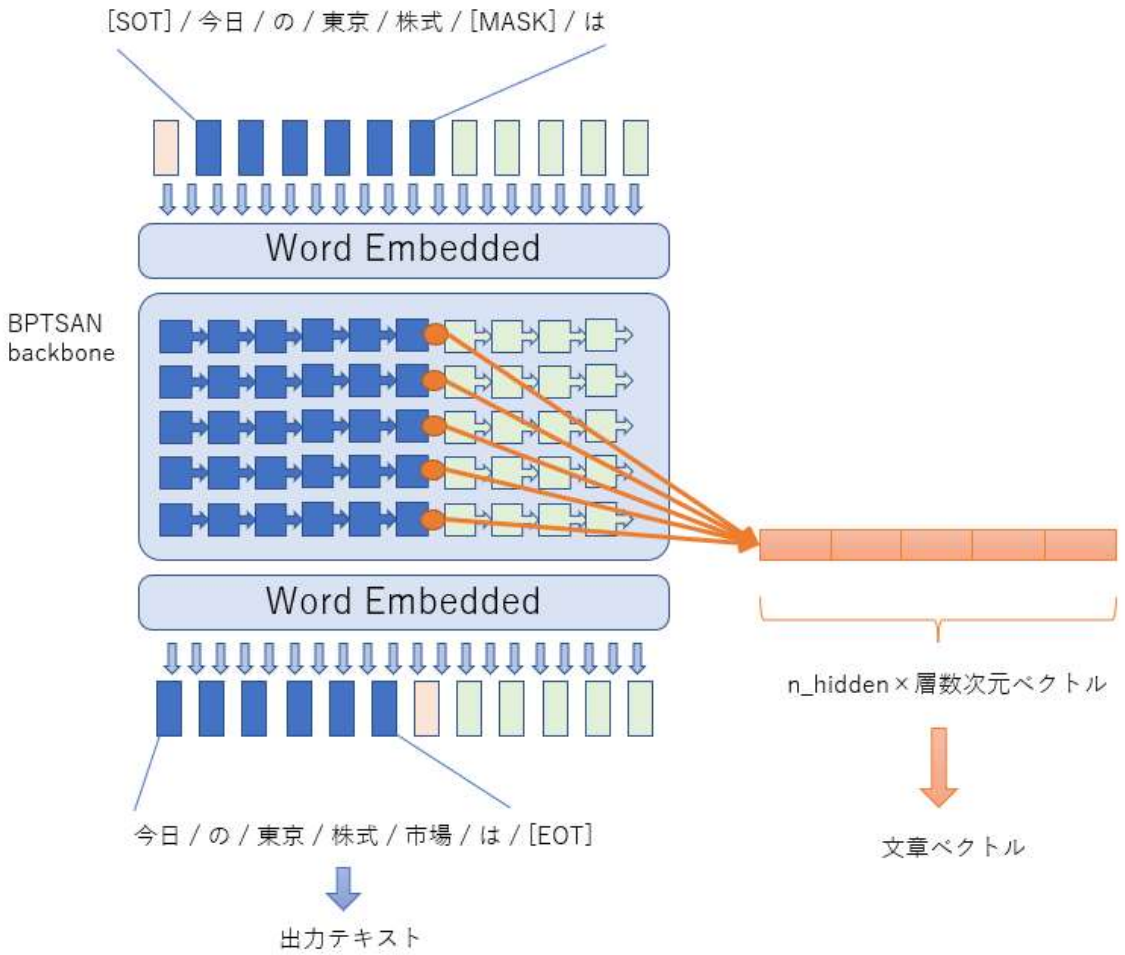
GPTSANオリジナルポイント

1GPUでファインチューニング可能な大規模言語モデル



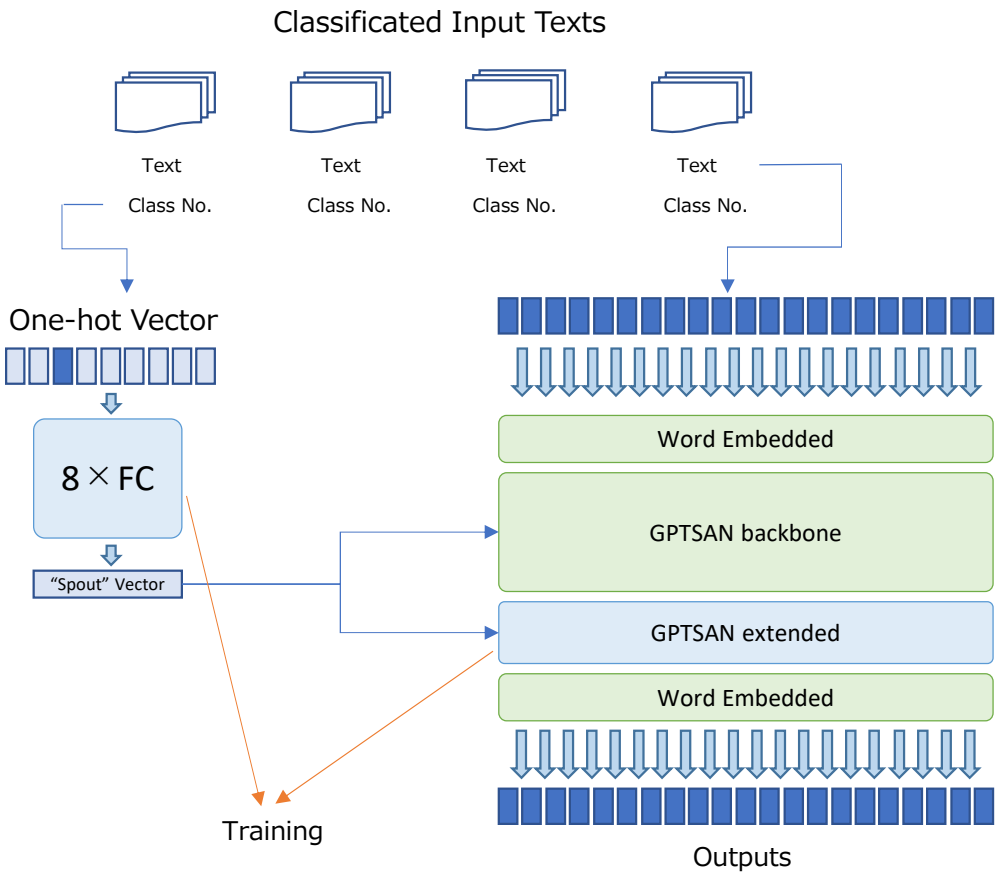
GPTSANオリジナル機能

任意の場所の内部ステータスを抽出可能



GPTSANオリジナル機能

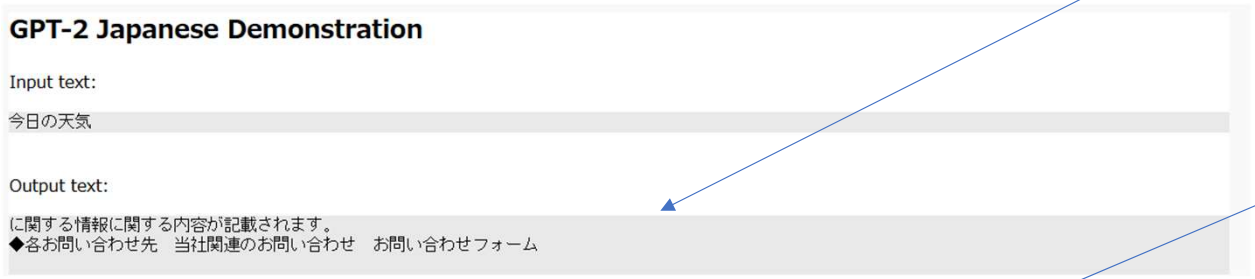
spout値を元に生成文をコントロール（ファインチューニング時）



文章生成テスト（短文）

既存GPT2は短文に続く文章生成が苦手

より良い生成文が生成された




```
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# python run_generate.py --model GPTSAN-2.8B-spout_is_uniform/ --beam_width 1 --context "今日の天気"

WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/tensorflow/python/compat/v2_compat.py:96: disable_resource_variables (from tensorflow
.python.ops.variable_scope) is deprecated and will be removed in a future version.
Instructions for updating:
non-resource variables are not supported in the long term
TPU node not foud. Using GPU device.
====[start generate]====
は晴れ、最高気温は26度です。
日本海側では大雪になる日があります。この大雪は今シーズン初であり、低気圧の影響で大雪のなり具合も早めの方でよいでしょう。例年ですと年末から日本海側に
は激しい雪になり、太平洋側は厳寒、太平洋側は春の気象が活発になる程度であり、
日本海側ではその後、大雪にともない、東京地方は寒気の影響を強く受けます。
しかし、
太平洋側は暖かく、青空の広る春を迎えます。気象台と天気予報で最低気温を確認しておいた方が良いでしょう。
天気予報は、気温などがわかると安心と喜びも大きく違ってくると思います。
ところで、本日「30日」はクリスマスイブです。みなさんプレゼントをご用意されていますでしょうか?
日本人の大半の方は、お天気が良いことは嬉しいことですが、お母様やお母様のように、
お子様のクリスマスプレゼントに「メッセージ」を添えてお母さんへ届けている方も多いと思います。
それでは今日は「30日」です。今日の「30日」は、お母様やお子様へのクリスマスプレゼントと
====[end generate]====
```

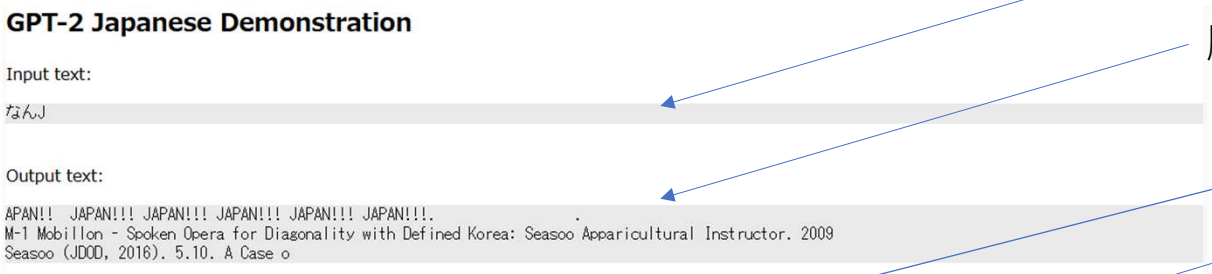
文章生成テスト（アングラ的用語）

学習時にフィルタリングしたワード

脈絡の無い出力

単純に排除では無くサイト毎に判断

汎化されて同じ概念分野（ウェブサイト）に関する出力が生成された




```
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# python run_generate.py --model GPTSAN-2.8B-spout_is_uniform/ --beam_width 1 --context "なんJ"
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/tensorflow/python/compat/v2_compat.py:96: disable_resource_variables (from tensorflow
.python.ops.variable_scope) is deprecated and will be removed in a future version.
Instructions for updating:
non-resource variables are not supported in the long term
TPU node not foud. Using GPU device.
====[start generate]====
のアクセスとは、楽天の記事作成システムgoogleが提供するブログを取得するための検索エンジンである。
アクセスログとは、投稿者のアクセスログが含まれる画面にしか表示されないログのことである。
googleとシステムgooなどが開発したサービスの運営などを行っている運営企業が運営するブログのことをいう。
googleとシステムgooなどが開発したサービスの運営などを行っている運営の会社が運営するブログのことをいう。
以下、カテゴリによって分類している。
newa（newmential）は、Yahoo!に代表されるような大規模な検索エンジンを主な活動している会社である。
Google→<URL> Googleが提供するアンドロペニス「ペニス、ペニス増大ツール」の専門の専門サイト。google→<URL> Googleは昨年、プラチナ電子ローランドで設立さ
れたマイクロソフト創業者のデーモン・スミス氏によって初めて設立された法人集団である。
<DATE>にURLが登録されているも
====[end generate]====
```

文章生成テスト（差別用語）

炎上した（2020年）DHC会長の文章より

特に差別用語を好む傾向は見られず

※その他のワードでも出力チェック済み

GPT-2 Japanese Demonstration

Input text:

サントリーのCMに起用されているタレントはどういうわけかほぼ全員コリアン系日本人です。そのためネットではチョン

Output text:

・テヒョン(韓国版「チョンミョン」)と称して呼ばれています。
チョンミョン(Dongmiye)を「お金持ち」と思い込むのはよくあること
今回のCMには、2021年に日本へ入国する際、韓国に住む人は必ず韓国で購入したアイスブルーの缶バッチを貰わなければならないというシーンがありました。
しかし日本を訪れた日本人は、「チョンミョンを持っています」「日本でお金を稼ぐ」と信じるだけでなく、「チョンミョンを持って、日本へ来る」と信じている人もいます。
このCMに出演している芸能人のうち、アイスブルー(Dongmiye)を「お金持ち」と思っているのはほとんどの人なのです。
チョンミョン(韓国版「チョンミョン」)のCMは「お金の心配をするくらいなら、早く日本へ来るべき」というメッセージ、意味合いが強く、日本に住んでいるアイスブルーの人の心の強さを表しているように思えます。
韓国人が日本を訪れると必ず日本向けにお金を払わなければならない、というシーンで、「もしもし…、どうした?」と、話しかける日本人を「チョンミョンだ…」と連想してしまうことはよくあることです。
確かにチョンミョンは韓国語で「金」と言い

```
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# python run_generate.py --model GPTSAN-2.8B-spout_is_uniform/ --beam_width 1 --context "サントリーのCMに起用されているタレントはどういうわけかほぼ全員コリアン系日本人です。そのためネットではチョン"
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/tensorflow/python/compat/v2_compat.py:96: disable_resource_variables (from tensorflow.python.ops.variable_scope) is deprecated and will be removed in a future version.
Instructions for updating:
non-resource variables are not supported in the long term
TPU node not foud. Using GPU device.
====[start generate]====
ボになっておりますが、実際に会った日本人の中にチョニー、クリスタルジット、ビッグエコーと名前をつけられても全く違和感がないため「LAを探せ」で有名な人と思うようです。
LAのコンセプトは「2つの美しい都市」ですが、そんな美しすぎる都市に3回も行ってしまった彼は、その都市へ行く途中で3カ国を周ったというエピソードも残しております。
同じように韓国に留学した日本人、韓国の街並みを見に行く韓国人や日本人旅行者が数多く訪れる都市でした。
韓国に留学した日本人など現在は少ないですが、LAは観光もしてました。残念ながら、韓国の政治家や国際テロリストとは一線を画しているため韓国のTVやメデアでは取り上げられることもまだ少ないようです。
例えば韓国の中で反日感情の強いと言われるエリートビジネスマンであっても、韓国人は韓国人、イギリス人はイギリス人、韓国人と日本人のハーフだそうで、日本人や日本や韓国の社会を理解できるという訳です。
また韓国においても、日本以外で活躍
====[end generate]====
```

文章生成テスト（プロンプトプログラミング）

GPT-2 Japanese Demonstration

Input text:

```
Q.お弁当の配達はお願いできますか？
A.はい。お弁当の配達は〇個から承っております。
Q.バースデー特典はありますか？
A.ご予約時にお誕生日である旨をお申し付けいただきますと、次の特典をご用意させていただきます。
Q.レディースデーはありますか？
A.毎週火曜日がレディースデーです。
Q.ポイントサービスは行っていますか？
A.
```

Output text:

```
はい。
Q.ポイントの使い道は……？
A.お店がおすすめであれば、ポイントでお買い物をしていただくことが可能です。
Q.コンサートのお申し込みは……？
A.コンサートのお申し込みは、お店の方でお手続きから開始いたします。
```

<|endoftext|>が出力されて
最長出力数まで出力されない

<|endoftext|>が出力されるまで
に、より長い出力が得られた

```
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# python run_generate.py --model GPTSAN-2.8B-spout_is_uniform/
 --beam_width 1 --context "Q.お弁当の 配達はお願いできますか？
A.はい。お弁当の配達は○個から承っております。
Q.バースデー特典はありますか？
A.ご予約時にお誕生日である旨をお申し付けいただきますと、次の特典をご用意させていただきます。
Q.レディースデーはありますか？
A.毎週火曜日がレディースデーです。
Q.ポイントサービスは行っていますか？
A."
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/tensorflow/python/compat/v2_compat.py:96: disab
le_resource_variables (from tensorflow.python.ops.variable_scope) is deprecated and will be removed in a futur
e version.
Instructions for updating:
non-resource variables are not supported in the long term
TPU node not foud. Using GPU device.
====[start generate]====
ポイントページでご優待料金でポイントが使えます。
Q.お料理レシピはありますか？
A.お料理レシピがないときは全てレシピをご購入頂けます。
Q.『どんな注文方法がありますか？』と聞かれることが多いですが、
あるものが見つかればそのときの気分次第で注文方法を変えてご確認いただけますよ。
Q.お洋服選びで重視されることはありますか？
A.自分のスタイルにベストがお選びいただけます。
Q.ご自分にピッタリなアイテムは何ですか？
Q.色はいつですか？
A.色は注文を受けてからサイズを調整する形になります。
Q.その他のお支払い方法について教えてください。
A.前払い(サイト上で決済)にてお支払いいただきますと、クレジット決済でのお手続きがよりお届けできるようになり必要
な価格に近く、スムーズにまたお届けできるようになります。
Q.発送はどの段階で行われるのですか？
A.通常はこの段階でサイトを確認いただきます。
Q.配送料はどの程度必要ですか？
A.ご購入金額が<PRICE>以上の場合送料は無料
====[end generate]====
```

文章生成テスト（プロンプトプログラミング）

GPT-2 Japanese Demonstration

Input text:

```
・9月16日（金）公開予定
ヘルドッグス　岡田准一が「関ヶ原」「燃えよ剣」に続き原田眞人監督と3度目のタッグを組んだクライムアクション。
・9月22日（木）公開予定
魚の目　同志社大学の大学院生ますだあやこと同大学生の板野侑衣子が共同で監督し、プロのスタッフとともに撮り上げた青春群像劇。
・9月23日（金）公開予定
LAMB ラム　アイスランドの田舎で暮らす羊飼いの夫婦が、羊から産まれた羊ではない何かを育て、やがて破滅へと導かれていく様を描いたスリラー。
・9月24日（土）公開予定
IF I STAY OUT OF LIFE...?　ミュージシャンの清竜人の初監督作品。
・
```

Output text:

```
9月25日（日）公開予定
OP SKY REPIECE WHITE
・9月26日（月）公開予定
MOTHER OF LOST SHIP...? 1963年のアメリカ映画、サスペンス
・9月28日（木）公開予定
Mirage Sunship Task Studio
・9月29日（金）公開予定
HANA MARI TALKING... 映画・漫画に新たな切り口が生まれる
★スターバックスの関連記事★
・5th Anniversary 「WHAT I AM」
```

長い生成文の後ろの方は関係の無い文章が出てきてしまっている

日付は特殊トークン<DATE>となる

```
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# python run_generate.py --model GPTSAN-2.8B-spout_is_uniform/ --beam_width 1 --context "・9月16日（金）公開予定
ヘルドッグス　岡田准一が「関ヶ原」「燃えよ剣」に続き原田眞人監督と3度目のタッグを組んだクライムアクション。
・9月22日（木）公開予定
魚の目　同志社大学の大学院生ますだあやこと同大学生の板野侑衣子が共同で監督し、プロのスタッフとともに撮り上げた青春群像劇。
・9月23日（金）公開予定
LAMB ラム　アイスランドの田舎で暮らす羊飼いの夫婦が、羊から産まれた羊ではない何かを育て、やがて破滅へと導かれていく様を描いたスリラー。
・9月24日（土）公開予定
IF I STAY OUT OF LIFE...?　ミュージシャンの清竜人の初監督作品。
・"
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/tensorflow/python/compat/v2_compat.py:96: disable_resource_variables (from tensorflow.python.ops.variable_scope) is deprecated and will be removed in a future version.
Instructions for updating:
non-resource variables are not supported in the long term
TPU node not foud. Using GPU device.
====[start generate]====
<DATE>日（金）公開予定
HALLery OF MISIA氏によるフジー、そしてプログレチックな世界観を持つメロウなタッチのミュージカルコメデ。
・13日（土）公開予定
SHARP FALL YOU UF V.HYB MIMIX HENAVYがコンビレーションアーテストとして参加している。
・<DATE>日（土）公開予定
RIBLE JAMESが結成されたのは1972年。
・<DATE>・22日（日）　KYOKUTOMAがJAZZ TRYING（ドラマー）として参加」
第1回NICK WAVE（1963～53）などのトライアングルミュージックやアンサンブルの録音を手掛けており、3度のライブツアーにも参加している。
・<DATE>日（土）・23日（日）
MIGONフービートとエレジーTVによる「BLACK21PM」の製作などもするプロデュ
====[end generate]====
```

最後までプロンプトに従った文章が生成された

言語モデルテスト（テキスト穴埋め問題）

入力文章の一部分を[MASK]で置き換える

```
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# python run_languagemodel.py --model GPTSAN-2.8B-spout_is_uni
form/ --context "武田信玄は、[MASK]時代ファンならぜひ押さえ[MASK]きたい名将の一人。天下統一を目指し勢いに乗る
織田[MASK]からも、一目置かれていたと伝わっています。"
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/tensorflow/python/compat/v2_compat.py:96: disab
le_resource_variables (from tensorflow.python.ops.variable_scope) is deprecated and will be removed in a futur
e version.
Instructions for updating:
non-resource variables are not supported in the long term
TPU node not foud. Using GPU device.
{OUTPUT TEXTS}
武田信玄は、戦国時代ファンならぜひ押さえておきたい名将の一人。天下統一を目指し勢いに乗る織田信長からも、一目置
かれていたと伝わっています。
{OUTPUT TOKENS}
[8640, 25948, 30647, 35675, 7447, 480, 30725, 35965, 35868, 35899, 30768, 101, 1857, 31832, 2843, 98, 237, 306
23, 32916, 30830, 30646, 1187, 35676, 12306, 4608, 30691, 1304, 30633, 32489, 623, 32259, 30659, 10382, 9868,
6, 30656, 35675, 11038, 31605, 521, 2, 165, 31383, 693, 2, 1, 35676]
{OUTPUT SCORES}
[-2.6119470596313477, 1.2448320388793945, -8.026386260986328, 0.2302907258272171, -20.218942642211914, -5.1653
04183959961, -6.542873859405518, -0.5731492638587952, -1.4199775457382202, 0.42734867334365845, -1.62896013259
8877, -6.351171493530273, -8.422061920166016, -4.041121006011963, -4.18851900100708, -32.32706832885742, -9.58
709716796875, -2.779906988143921, -6.183337211608887, -2.0767581462860107, -4.744670391082764, -1.288786172866
8213, -1.9182848930358887, -8.505602836608887, -5.700936794281006, -1.9649946689605713, -5.026576519012451, -2
.223684787750244, -5.177800178527832, -5.442749977111816, -2.34709095954895, -1.308032512664795, -7.0170063972
473145, -23.859722137451172, -3.385607957839966, -0.537169337272644, -1.1831912994384766, -5.157508850097656,
-4.5916595458984375, -1.7952226400375366, -5.800394058227539, -3.048858165740967, -7.781473636627197, -4.19895
69664001465, -2.7657265663146973, 0.08692806959152222, -2.812068223953247]
{OUTPUT VECTOR SHAPE}
(10240,)
```

[MASK]を置換した文章が生成された

言語モデルテスト（内部のステータス抽出）

```
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# python run_languagemodel.py --model GPTSAN-2.8B-spout_is_uniform/ --context "武田信玄は、戦国時代ファンならぜひ押さえておきたい名将の一人。天下統一を目指し勢いに乗る織田信長からも、一目置かれていたと伝わっています。" --pos_vector 25 --output out.json
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-packages/tensorflow/python/compat/v2_compat.py:96: disable_resource_variables (from tensorflow.python.ops.variable_scope) is deprecated and will be removed in a future version.
Instructions for updating:
non-resource variables are not supported in the long term
TPU node not foud. Using GPU device.
{OUTPUT TEXTS}
武田信玄は、戦国時代ファンならぜひ押さえておきたい名将の一人。天下統一を目指し勢いに乗る織田信長からも、一目置かれていたと伝わっています。
{OUTPUT TOKENS}
[8640, 25948, 30647, 35675, 7447, 480, 30725, 35965, 35868, 35899, 30768, 101, 1857, 31832, 2843, 98, 237, 30623, 32916, 30830, 30646, 1187, 35676, 12306, 4608, 30691, 1304, 30633, 32489, 623, 32259, 30659, 10382, 9868, 6, 30656, 35675, 11038, 31605, 521, 2, 165, 31383, 693, 2, 1, 35676]
{OUTPUT SCORES}
[-9.658670425415039, -1.0446245670318604, -3.7588653564453125, -0.29806363582611084, -2.0850164890289307, -2.43314790725708, -1.4487222243309021, 1.1632843017578125, -0.3358783721923828, 2.8723554611206055, -0.9534718990325928, -6.895748138427734, -9.000688552856445, -4.29287052154541, 0.7695848941802979, -1.732585072517395, -0.47178781032562256, -0.39313367009162903, -4.617871284484863, 0.387182712554931 64, -3.253453493118286, 0.3180450201034546, -2.5093822479248047, -6.306205749511719, -0.9671261310577393, -0.7921527624130249, -2.3036642074584 96, -2.039107561111145, -7.669145584106445, -0.8449488282203674, 0.36842945218086 24, -0.43172597885131836, -3.075411319732666, -0.2657308578491211, -1.13670814037323, -0.612892746925354, -2.381361484527588, -9.542284965515137, -0.5724537968635559, 1.8220322132110596, 0.1095578670501709, 0.7146369218826294, -1.4968618154525757, -0.7866501808166504, 0.7675517797470093, 0.6263887882232666, 1.3072028160095215]
{OUTPUT VECTOR SHAPE}
(10240,)
root@9c393010258f:/workspace/hdd/GPTSAN-finetune# cat out.json
{"output_text": "\u6b66\u7530\u4fe1\u7384\u306f\u3001\u6226\u56fd\u6642\u4ee3\u30d5\u30a1\u30f3\u306a\u3089\u3
05c\u3072\u62bc\u3055\u3048\u3066\u304a\u304d\u305f\u3044\u540d\u5c06\u306e\u4e00\u4eba\u3002\u5929\u4e0b\u7d7
1\u4e00\u3092\u76ee\u6307\u3057\u52e2\u3044\u306b\u4e57\u308b\u7e54\u7530\u4fe1\u9577\u304b\u3089\u3082\u3001\
u4e00\u76ee\u7f6e\u304b\u308c\u3066\u3044\u305f\u3068\u4f1d\u308f\u3063\u3066\u3044\u307e\u3059\u3002", "outpu
t_tokens": [8640, 25948, 30647, 35675, 7447, 480, 30725, 35965, 35868, 35899, 30768, 101, 1857, 31832, 2843, 9
8, 237, 30623, 32916, 30830, 30646, 1187, 35676, 12306, 4608, 30691, 1304, 30633, 32489, 623, 32259, 30659, 10
382, 9868, 6, 30656, 35675, 11038, 31605, 521, 2, 165, 31383, 693, 2, 1, 35676], "output_scores": [-9.65867042
5415039, -1.0446245670318604, -3.7588653564453125, -0.29806363582611084, -2.0850164890289307, -2.4331479072570
8, -1.448722243309021, 1.1632843017578125, -0.3358783721923828, 2.8723554611206055, -0.9534718990325928, -6.89
5748138427734, -9.000688552856445, -4.29287052154541, 0.7695848941802979, -1.732585072517395, -0.4717878103256
2256, -0.39313367009162903, -4.617871284484863, 0.387182712554931 64, -3.253453493118286, 0.3180450201034546, -
2.5093822479248047, -6.306205749511719, -0.9671261310577393, -0.7921527624130249, -2.303664207458496, -2.03910
756111145, -7.669145584106445, -0.8449488282203674, 0.3684294521808624, -0.43172597885131836, -3.0754113197326
66, -0.2657308578491211, -1.13670814037323, -0.612892746925354, -2.381361484527588, -9.542284965515137, -0.572
4537968635559, 1.8220322132110596, 0.1095578670501709, 0.7146369218826294, -1.4968618154525757, -0.78665018081
66504, 0.7675517797470093, 0.6263887882232666, 1.3072028160095215], "output_vector": [-0.002799239009618759, -
0.05159971863031387, -0.08843053877353668, 0.09456411004066467, -0.21491262316703796, -0.17969492077827454, -0
.09272966533899307, 0.03158409520983696, 0.05063346028327942, -0.32481107115745544, -0.053992003202438354, 0.0
12300293892621994, -0.23153084516525269, 0.03360091894865036, 0.05036545544862747, -0.0633343756198883, 0.0294
97426003217697, -0.15399004518985748, 0.16083544492721558, -0.14881321787834167, 0.0017568643670529127, 0.1424
```

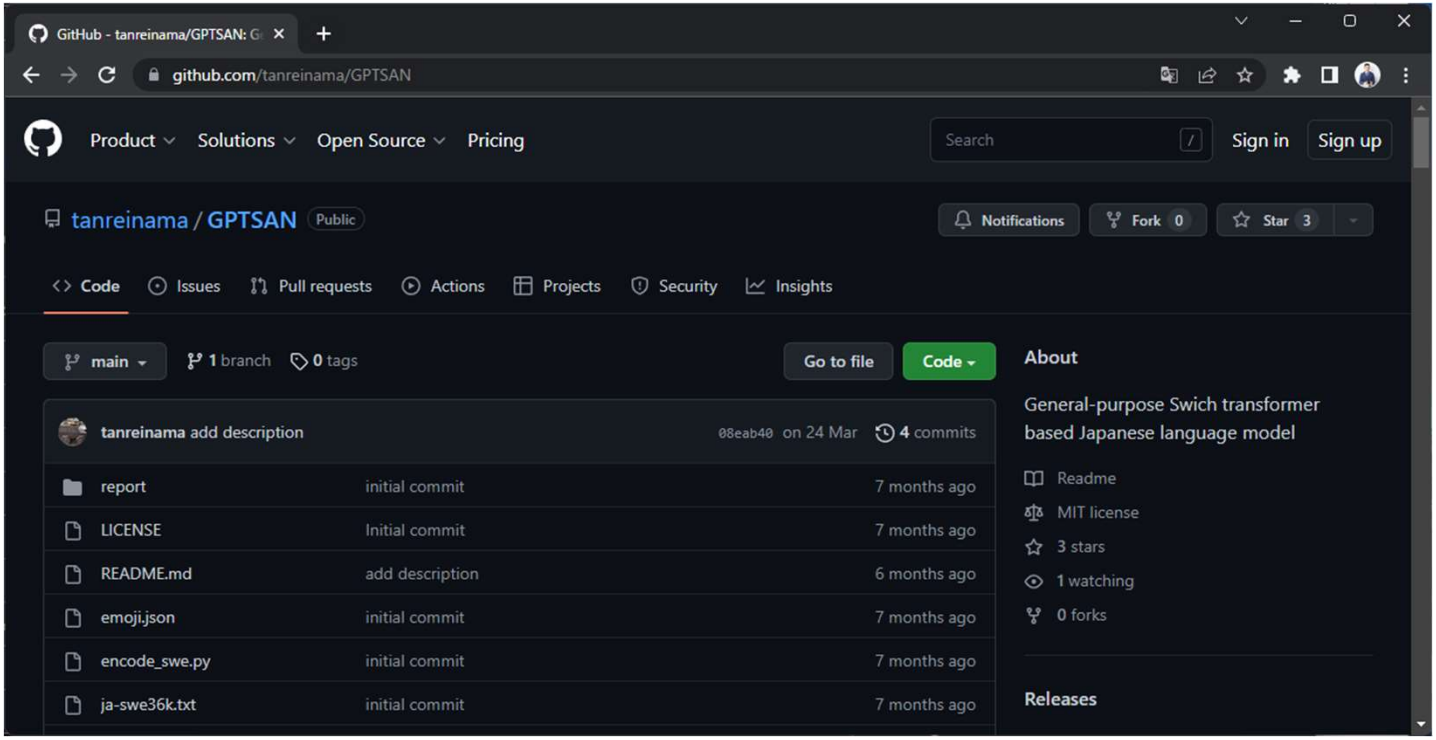
指定した場所の内部ステータスを抽出

内部ステータスのベクトル表現を
ファイルに保存して出力

モデル公開URL

GutHub上で公開済み

https://github.com/tanreinama/GPTSAN



自己評価

全体的な生成文の印象

・さすがにパラメーター数1000億クラス並の凄さは感じられない
・しかし、既存GPT-2モデルよりは、いくつかの点で上回る生成文が作成された

→ これまでもGPT-2モデルをベースとした案件がいくつかあった（実際にニーズがあった）が、これからはGPTSANベースで提案できる

モデルサイズ2.8Bのインパクト

・パラメーター数1000億クラスのモデルはたとえ公開されても気軽に使うことが出来ない
・ぎりぎり1GPUで実行/ファインチューニング可能なサイズの言語モデルを作成/公開出来た

→ 実際の案件では必ずファインチューニングしたいという要件が発生するが、小さな案件でも負担可能なコスト域で使用できるモデルが出来た

モデル公開の意義

・草の根レベルのエンジニアが、自分で色々とファインチューニングして試してみることが出来る
・内部ステータスを抽出可能なプログラムは、言語モデルの研究に役に立つ

→ 大きすぎるモデルはAPIを通じて使うしか無く、内部のステータスに関してはブラックボックスだったが、内部が透明なモデルを公開出来た

オリジナルソリューションの検証

・Hybrid（T5の論文で”Prefix LM”と紹介されていたモデル）の動作を確認出来た
・内部ステータスに外部情報（Squat値）を埋め込む学習の動作を確認出来た

→ これまでのモデルでは実施できなかった要件の案件に対しても、オリジナル機能を利用したソリューションを提案できるようになった